

**FIGURE 1**

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTTCAGATCTGCTCGGTAGA  
CCTGGTGCACCACCACC**ATG**TTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG  
GTTTTCCACCCAGCTTTTACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA  
ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAACAAGAATTGGGATCCGGCGTGGGA  
GAACTGGCCAAGAAGCTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT  
GATCAGATGGGAAGATGGTTTGTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA  
CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT  
ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAAACAGCT  
TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACCTTCATGATGAGAGGCTCTTG  
GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC  
CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT  
GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTTCTCATCAGAGCTGCATGGTACAC  
AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCAGTGAAAAGTTTCTGA  
ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG  
TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT  
AGTTCTTTTTCAGCATGTTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT  
CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACCTCGATGCTGAGTATCTACATGGAT  
ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAA**TG**  
**A**AGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAAATGGGGCAGATATGC  
ATTAAATAGTTTGTACAAGCAGCTTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA  
TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT  
AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT  
TTGGTGAATGTGAAAATAAGTTTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA  
TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT  
GCAGAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGGACCAGAGAGAAG  
GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA  
GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTCATTT  
GCTGAACTTAACAAAAGTTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG  
CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC  
ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTACTTTTGAATGTTACAAAAGGAA  
ATAACTTTAAAAGTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG  
AATACAAACAGTATACTCATG

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL  
KEAALEPSMEKIFKIDQMGRWVAGGAAGVLGALCYYGGLSNEIGAIEKAVIWPQYVKDRI  
HSTMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP  
GPKHLAWLLHSGVMGAVVAPLTLGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL  
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFMSMFLLYDTQKVIKRAEVSPMYGV  
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

**FIGURE 3**

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCGTCTCCGCCTTCTGCAT  
 CGCGGCTTCGGCGGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGGCCGCTCGTGAGGG  
 GGTCCGCACGGGGAGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTCTGAAG**ATG**TCGG  
 ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCCGCCCCACCGTC  
 GCCGTGCCCTTGGTCGGCAAACCTCGGCCTCATCAGCCCGGCCTACCTCTTCTCTGGCCCGA  
 AGCCTTCCCTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTTATTTCCCTGTGG  
 GTCCAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA  
 CTTGAAACAGGAGCTTTTGTGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTTAACTG  
 GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCCTCTGATCA  
 TGTCAGTACTTTATGTCTGGGCCAGCTGAACAGAGACATGATTGTATCATTTTTGGTTTGGGA  
 ACACGATTTAAGGCCTGCTATTTACCTGGGTATCCTTGGATTCAACTATATCATCGGAGG  
 CTCGGTAATCAATGAGCTTATTTGAAATCTGGTTGGACATCTTTATTTTTTCTAATGTTCA  
 GATACCCAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGTGG  
 CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC  
 TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC  
 AG**TGA**AGGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCTCCAGTGCTGGGTG  
 CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC  
 AGTACGAGACAAAGTTTCTTAAATCCGAAGAAAAATATAAGTGTTCCACAAGTTTACGAT  
 TCTCATTCAAGTCCTTACTGCTGTGAAGAACAATAACCAACTGTGCAAATTGCAAAACGAC  
 TACATTTTTTGGTGTCTTCTCTTCTCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCCT  
 AATCTGCTGGCATTGAGCTGGGGCTGGGTCAACCAACCCTTCCCAAAGGACCTTATCTCTT  
 TCTTGCACACATGCCTCTCTCCCACTTTTCCCAACCCCCACATTTGCAACTAGAAAAAGTTG  
 CCCATAAAATTTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGT  
 ACAACAATCATATTCACGTTATTTTCCCCTTTTGGTGGCAGAAGTGTACCAATAGGGGGAG  
 AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATTGCTTCGACTGACATCCGTTGTT  
 AACCGTTTGGCACTCTTCAGATATTTTTTATAAAAAAGTACCACTGAGTTCATGAGGGCCA  
 CAGATTGGTTATTAATGAGATACGAGGGTGGTGTGGGTGTTTGTTCCTGAGCTAAGTGA  
 TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCC  
 TTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTGCTTTGTGT  
 TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCTGGGCATTG  
 ATTCCCATTTCATCTCATTCTGGATATGTGTTCAATTGAGTAAAGGAGGAGAGACCCTCATA  
 CGCTATTTAAATGTCACTTTTTTGCTATCCCCCGTTTTTGGTTCATGTTTCAATTAATTGT  
 GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTAAAGCTAATGTAAGCACATCTA  
 AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA  
 TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT  
 TCGTAGGTGGGCTTTTCCCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA  
 TGGCTTTTTCACACAGTTATTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT  
 ATTGAGTGGCTGTACACACTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTGATCAGTTTCTT  
 TTCAGGAAACATTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT  
 GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG  
 TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAAGTGTCA  
 TGTTTTGTGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAAAATTTTAGAGGCCATTACT  
 CCAATTATGTTGCACGTACACTCATTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA  
 TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCCTCTTACCAGTCAGCTGCCTGCGAG  
 CAGTCATTTTTTCTAAAGGTTTACAAGTATTTAGAACTTTTCAGTTCAGGGCAAAATGTTT  
 ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT  
 ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA  
 TGACAGTGGATTCTCTTTACAAATGGAAAAAAAATCCTTATTTTGTATAAAGGACTTCCC  
 TTTTTGTAAACTAATCCTTTTTATTGGTAAAAATTGTAAATTAAATGTGCAACTTG

**FIGURE 4**

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF  
 PVGPGTGFLYLVLNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP  
 LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIIGGSVINELIGNLVGHLYFFL  
 MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL  
 GDQ

**Transmembrane domain:**

amino acids 98-116, 152-172

**N-myristoylation site.**

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

**Glycosaminoglycan attachment site.**

amino acids 218-222

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF  
 PVGPGTGFLYLVLNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP  
 LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIIGGSVINELIGNLVGHLYFFL  
 MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL  
 GDQ





**FIGURE 6**

MAAEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP  
LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEAactsPLAKTHTSQAILQP  
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVVSdleHEEMKILREVL  
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS  
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPiANLSVLGTEELRQREHYLKQKRDKLMS  
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEKQTLLKRLLAEKLKEEVINK

**N-glycosylation sites.**

amino acids 224-228, 246-250, 285-289

**N-myristoylation site.**

amino acids 273-279

**Amidation site.**

amino acids 252-256

**Cytosolic fatty-acid binding proteins.**

amino acids 78-108





**FIGURE 9**

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK  
 LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLPAMAV  
 IFSNFSIITALLFRIVLKRRNLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDFAFF  
 SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNNTARVFSHIRLGMGHVLIIVQCFISSMANI  
 YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI  
 FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI  
 YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDETF

**Transmembrane domains:**

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,  
 298-318, 328-368

**N-glycosylation sites.**

amino acids 128-132, 204-208, 218-222, 374-378

**Glycosaminoglycan attachment site.**

amino acids 402-406

**N-myristoylation sites.**

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK  
 LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLPAMAV  
 IFSNFSIITALLFRIVLKRRNLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDFAFF  
 SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNNTARVFSHIRLGMGHVLIIVQCFISSMANI  
 YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI  
 FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI  
 YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDETF

CGTGCCTGCGCAATGGGTGTCTGGGTCCGCTTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG  
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTTCGCTATACCTACTGTAGCTTCTCCAC  
GTATGGACCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG  
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAACAAACAGTGGAATGGAA  
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATACATTCTGCTAGG  
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG  
AAAACAAGTATGATTATCTTCCAACACTACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT  
TTCTGTGTGCTTGTGTGCTTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA  
TGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCTGCTTTCTTTATTTCC  
TGGATAACTTGATTGTCTTCTATGTCTGCTCTATCTTCAACCAGCCATGGCTGTTATCTTC  
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAAGAGGCGTCTAAA  
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA  
CTAAAACCTTTA

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGCCGGCTTGCTAGCGCGCGGGCCGTG  
GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCTGCGGGGCGAGAGGAGCAT  
CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGGGGTTCATGGCCAAAGGAGAAGGCGCCGAG  
AGCGGCTCCGCGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCA  
GGTGAAGAAAGAACC GAAAAAGAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG  
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCTTCAGATCTAC  
CTATTG**GATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC  
CTGGGATGCCATCACAGACCCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC  
TGGGTGCGCTTATGCCCTGGATCATCTTCTCCACGCCCCCTGGCCGTCAATTGCCTACTTCCCT  
ATCTGGTTCTGTGCCGACTTCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT  
CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA  
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC  
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCCTTGTTCAGG  
ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATAACATGGCACCACCTTCACAC  
AGGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCAATTGTCTGTATCTATATAATCTG  
TGCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCAGCAGTCTG  
AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCT  
ATTACTGGCTTCTCTTACCTCCTTGGCTTTTCATGCTGGTGGAGGGGAACCTTTGTCTTGT  
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT  
CGGCCACTTTAACCATTCCCCTCTGGCAGTGGTTCTTGACCCGGTTTGGAAGAAGACAGCT  
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA  
CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC  
TACCCTGGTCCATGCTGCCTGATGTCAATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT  
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTCACCAAGTTTGCCTCTGGAGTGT  
ACTGGGCATTTCTACCCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC  
CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG  
CTGGGCCTGCTGCTCTTCAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAATAAGAA  
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG  
AGCTGGCTAGCATCCTC**TAG**GGCCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG  
AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA  
CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG  
TGGCCTCCTGCCTCCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA  
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTACAGAGCT  
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA  
TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

[illegible][illegible]



**FIGURE 13**

GGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT  
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA  
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA  
TACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT  
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC  
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG  
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC  
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT  
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG  
GAACCGAGCCCAT

**FIGURE 14**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT  
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATGT**  
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT  
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT  
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG  
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA  
GAGAACGTTATCATCAAATTAACAAGGCTGGCCTTGTAAGTGGAGCTGAGTTGTTTAGG  
ACTTTCTATTGTGGCAAACCTTCAGAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG  
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG  
CAGCCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG  
AGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG  
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTATGTGCTTCACATGATCACT  
ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGA  
TTTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG  
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATT**TGA**TGAAAGGAT  
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAAGTTGCTTA  
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA  
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA  
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACTATG

MWWFQQGLSFLPSALVIWTSAAAFIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI  
AAVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG  
AVLTFGMGSLYMFVQTILSYQM QPKIHGKQVFWIRLLLLVIWCGVSALSMLTCSSVLHSGNFG  
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLT YIRDFQKISLRVEANLHGLTLYD  
TAPCPINNERTRLISRDI

**FIGURE 16**

CGGACGCTTGGGCNCGGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT  
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA  
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT  
TACCAAGAATCTTCAACCCTTTCCACAAAAGCTAATTGAGTACACGTTCCCTGTTGAGTACA  
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA  
GTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT  
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT  
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

CCACACGCGTCCGCCGCCGCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCGCGGG  
CCGGGGTGCGGAGCCGACATGCGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC  
CTTCGCCTTGTACTTGCTGTGCGACGCGACTGCCCGCGGGCGGAGACTGGGCTCCACCGAGG  
AGGCTGGAGGCAGGTGCGTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG  
GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCTTGCTCTTCTGCGGCGCCTA  
CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT  
TGTTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC  
TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCCTGATAAAGT  
GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTTCTTATTGTTTT  
TGAGACTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT  
CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTTCATCTGTGT  
GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG  
TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT  
AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA  
CACTGATCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA  
TGTGGTCCTCTAAAGCCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG  
CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT  
TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT  
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATT  
ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAAT  
ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC  
AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT  
CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

**FIGURE 18**

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRELSEVLREYR  
 KEHQAYVFLFLFCGAYLYKQGFaipgssflnvlagalfgpwlglllccvltsvgatccyllss  
 IFGKQLVVSYPDKVALLQRKVEENRNSLFFFLFLRLFPMPNWFNLNSAPILNIPIVQFF  
 FSVLIGLIPYNFICVQTGSILSTLTSLDALESWDTVFKLLAIAMVALIPGTLIKKFSSQKHLQ  
 LNETSTANHIHSRKDT

**Important features:****Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 101-123, 189-211

**N-glycosylation sites.**

amino acids 172-176, 250-254

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 240-244, 261-265

**N-myristoylation site.**

amino acids 13-19, 104-110, 115-121, 204-210

**Amidation site.**

amino acids 27-31

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 4-15

**Protein splicing proteins.**

amino acids 25-31

**Sugar transport proteins.**

amino acids 162-172

CCGAGGCGGGAGGAGCCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTT  
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACAGCTCAGAA  
TAGGAAAATAACTTGGGATTTTATATTGGAAGAC**ATG**GATCTTGCTGCCAACGAGATCAGCA  
TTTATGACAAACTTTCAGAGACTGTTGATTTGGTGAGACAGACCGCCATCAGTGTGGCATG  
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC  
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT  
TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT  
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA  
GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC  
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTGCTG  
AACTGCACTGGCTGTGCCGAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA  
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC  
AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTTCGCCAAG  
TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTTCTTATCCATGGAGGAGACCTCTGAA  
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCACTCACCTGCCATTTCCAAAAGATG  
CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG  
ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA  
GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCTG  
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT  
GGAACCGCTTTCTCAGAACTG**TAG**GAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA  
AAACCAGGTTGAAAGGGGAAAAATAAAAAACAAAACGATGAAACTGCAAAAA

**FIGURE 20**

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY  
KVLATLGLILLTAYFVIQPFSPPLAPEPVLGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL  
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMILLEDAPRKFERLHPLVIKT  
GKPLLEEEIQHFLCQYPEATEGEFSEGGFAKWWRCFPERWFPPYPWRRPLNRSQMLRELFV  
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRHCQSVAMP  
IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL



CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG  
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTCAGTTGTGATC  
AAGGGACACGTGGTTTCCGAAC TGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT  
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAAACTTTCAGAGACTGTTG  
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAGAGAAGGCAATTGAAAAATTTATC  
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCGCAGTATCCTCTCCTTATAGT  
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC  
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

CCACACGCGTCCGCCACACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG  
CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTG  
CCACTGGTGCGCAGCGTGTAGACCGTGCCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC  
TCTCCTGCCACCCACCAATGAGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA  
CACCCGAATGGCGCCACTTCATCGACAAACAGGTACAGCCAAACC**ATG**TCCAGTTCGAAATG  
GACACGTATGCTAAGAGCCACGACCTTATGTCAGGTTTCTGGAATGCCTGCTATGACATGCT  
TATGAGCAGTGGGCAGCGGCGCCAGTGGGAGCGCGCCAGAGTCGTCGGGCCCTTCCAGGAGC  
TGGTGCTGGAACCTGCGCAGAGGCGGGCGCGCCTGGAGGGGCTACGCTACACGGCAGTGCTG  
AAGCAGCAGGCAACGCAGCAGCTCCATGGCCCTGCTGCACTGGGGGGCGCTGTGGCGCCAGCT  
CGCCAGCCCATGTGGGGCCTGGGCGCTGAGGGACACTCCCATCCCCGCTGGAACATGTCCA  
GCGCCGAGACATATTACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTTCGACCCCTCAC  
CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCTGACACCCACCGAGGAGGC  
CTCACTGCCTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCACCCACCCGAGTTGCTGCAGG  
AGGACCACTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAAGCTG  
GATGAGCAGCGTGAGAAGCTGGTGCTGTCGGCCGAGTGCCAGCTGGTGACGGTAGTGCCCGT  
GGTCCCAGGGCTGCTGGAGGTCACCACACGAAGTGTATACTTCTACGATGGCAGCACTGAGC  
GCGTGGAACCCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCCAGCTGCGTGAG  
GTCCACCTGCGGCGTTTCAACCTGCGCCGTTTACGCACTTGAGCTCTTCTTTATCGATCAGGC  
CAACTACTTCTTCAACTTCCCATGCAAGGTGGGCACGACCCAGCTTCATCTCCTAGCCAGA  
CTCCGAGACCCAGCTTGGCCCCATCCCACCCATACCCAGGTACGGAACCAAGGTGTAATCG  
TGGCTCCTGCGCCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT  
GCTGCGTGCCCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACCTTCGAGTACT  
TGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTT  
CCCTGGGTCTCTGCAGGACTACGTGTCCCCAACCCCTGGACCTCAGCAACCCAGCCGTCTTCCG  
GGACCTGTCTAAGCCCATCGGTGTGGTGAAACCCCAAGCATGCCAGCTCGTGAGGGGAGAAGT  
ATGAAAGCTTTGAGGACCCAGCAGGGACCAATTGACAAGTTCACCTATGGCACCCACTACTCC  
AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGGAGCCCTTACCTCCCTGCACGTCCA  
GCTGCAAAGTGCGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCCTGGC  
AGGCACGCCTGGAGAGCCCTGCCGATGTGAAGGAGCTCATCCCGGAATTCTTCTACTTTCTCT  
GACTTCTGGAGAACCAACGAGTTTGTACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGT  
AGGCATGTGGTGCTACCCCCGTGGGCCAGCTCTCTGAGGACTTCATCCAGCAGCAGCCGCC  
AGGCTCTGGAGTCGGAGTATGTGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGC  
TACAAGCAGCGGGGGCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA  
GGGGCTCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGGCATTA  
TCAGCAACTTTGGGCAGACTCCCTGTCAAGCTGCTGAAGGAGCCACATCCAACCTCGGCTCTCA  
GCTGAGGAAGCAGCCCATCGCCTTGCACGCGCTGGACACTAAGTCACTTACGATCTTTCAGCA  
CCTGGACGAACTCAAGGCATTCTTTCGACAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA  
CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTTCAGCTTCAGCAAAGACCCC  
ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCCGTGGGTGCCAGGCAGTGGTGT  
GAGTGGAACAAGCACTGGCAGTGGCCCGGATGGAAAGCTGCTATTACGCGGTGGCCACTGGG  
ATGGCAGCCTGCGGGTGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC  
CTTGATGTAGTAACCTGCCTTGCCTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG  
GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCAC  
CAAAGCCTGTGCAGGTCCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT  
GAACCTGACATGGCTGTGTCTGATCTGAGGATGGAAGCTGTATCATACACACTGTACGCCG  
CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCTTGAGACCTATTTTCCACC  
TGGCATTTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCCTGGGGCC  
CAGGTACCTACTCCTTGCACCTGTATTAGTCAATGGGAAGTTGCGGGGCTTCACTGCCCCCT  
GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCAGT  
GCGCCCTGCACATCTCCAACATAACACTGCTCCCGCCGCGCTCCTTGGCCCATGAAG  
GTGGCCATCCGACAGCTGGCCGTGACCAAGGAGCGCAGCCAGCTGCTGGTGGCGCTGGAGGA  
TGGCAAGCTCATCGTGGTGGTTCGCGGGGCAGCCCTCTGAGGTGCGCAGCAGCCAGTTTCGCGC  
GGAAGCTGTGGCGGTCTCTCGCGGCGCATCTCCAGGTGTCTCGGGAGAGACGGAATACAAC  
CCTACTGAGGCGCGCT**TGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCCCGGCAGGCCTG  
GCCCGGAGGGCCCCGCGCAAGTCCGGCGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGA  
GCGGGGCCACCCCTGCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGAATTGGCG  
GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGGCCGCCCTGAGGGGCCAGCACTGGCGTCT

**FIGURE 23**

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL  
 RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN  
 HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAEELETP  
 MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP  
 LAQLREVHLRFRNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV  
 RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGR TYNDL  
 SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH  
 YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADV KELIP  
 EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPWASSPEDFIQQHRQALESEYVSAHLHEW  
 IDLI FG YKQRGPAEEALNVFYYCTYEGAVDL DHVTDERERKALEGIISNFGQTPCQLLKEP  
 HPTRL SAEAAHRLARLD TNSPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF  
 SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGSLRVTALPRGKLL  
 SQLSCHLDVVTCLALDTCGIY LISGSRDTTCMVWRL LHQGGLSVGLAPKPVQVLYGHGA AVS  
 CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA  
 WERPGAQVTYSLHLYSVNGKLRLASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA  
 PPLPMKVAIRSVAVTKERSHVLVGLLEDGKLIVVAGQPSEVRSSQFARKLWRSSRRISQVSS  
 GETEYNPTEAR

**N-glycosylation site.**

amino acids 677-681

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 985-989

**Tyrosine kinase phosphorylation site.**

amino acids 56-65, 367-376, 543-551

**N-myristoylation site.**

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,  
 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,  
 873-879, 912-918, 954-960

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC  
CACGGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT  
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC  
TGGACCTTAAGTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT  
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC  
GCACACTCCGTTACCACACTGGGTCATTGGCATTTGGAGCCCTCATCCTGACCCTTGTGCAG  
ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGC  
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC  
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA  
AATGCGTTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTTCCTGGACAAAGTCACAGA  
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCCTGTCCTTCTTTT  
TTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCACCTCAACTATTAC  
TGGCTGCCCATCATGACCTCCATCCTGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT  
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG  
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC  
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGC  
ACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT  
AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG  
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC  
GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC  
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA  
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA  
AAAGATTTTATTAAAGATATTTTGTTAAGTCT

[illegible]

**FIGURE 25**

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLG  
LFWTLNWWLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ  
IARVILEYIDHKL RGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK  
NAFMLLMRNIVRVVVLDDKVTDLFFFGLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY  
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLLKILGKKN  
EAPPDNKKRKK

**FIGURE 26**

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT  
 GGCT**ATG**TTCGTGTCCGATTTCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC  
 TTCTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTT  
 CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC  
 ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG  
 ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG  
 CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA  
 TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT  
 CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA  
 GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT  
 CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG  
 CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC  
 CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGCTCTGCAGCG  
 CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA  
 CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC  
 AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA  
 GCGGCTCCAGGAGTTCCTTGACAGCATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAAGTTCC  
 AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA  
 TTTGGGATGAAGGACATGCGCGTGCAGACTTTCAGCATTCAATTTGGGTTCAGGACAAAGTT  
 TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT  
 CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG  
 TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC  
 CTTTGACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC  
 TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA  
 AGTCCTTTGTGTGTTGACAAAGAACCGGCGCTGCAAACCTGCTGCCCCCTGGTGATGGCTGCC  
 CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC  
 GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA  
 TGCTGCACAACCATTTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT  
 CTGGACGCACTTATTTCCCTCCTGTCC**TAG**GAATTTGATTCTTCCAGAATGACCTTCTTATT  
 TATGTAACCTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT  
 TTTTATTAAATAAAATGCTTATTTTAGGAAA

**FIGURE 27**

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF  
LEHKEQFHYFILINCGANVDLLDILQPEDDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD  
LEVPAIEDIFRDEEEDEEHSGNDSGDSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF  
DYEQYEHGTSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLRH  
VSRHNRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR  
LQEF LADMGLPLKQVKQKFQAMD ISLKENLREMIEESANKFGMKDMRVQTFSIHFGFKHKFL  
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSNLDKLYHGLELAKKQLRATQQTIASCL  
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP  
LSMEHGTVTTVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNFHDL SVIELKAEDRSKFL  
DALISLLS

**FIGURE 28**

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGTGGCTATGNTCGTGTCCGATTTCCGCA  
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCTTCTCTTCGTGGCCTCGGGANGTGGA  
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT  
GGTTCAGTCTTCTGGGTGGCAAGAACTTGAAACTGCATTTCTTGAGCATAAAGAACAGTTTC  
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT  
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA  
CGATACCC





**FIGURE 30**

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPH  
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF  
FAFSVGRKTALHSGEDFQTLLFERVFNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET  
YVHIMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT  
FSGHLIKAEDD

**Important features:****Signal peptide:**

amino acids 1-20

**N-glycosylation site.**

amino acids 72-75

**C1q domain proteins.**

amino acids 144-178, 78-111 and 84-117



**FIGURE 32**

MCSRVP LLLPL LLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTPRDPVPPDTVGLYVFEN  
GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR  
RLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP  
GILDTANVEALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAG  
NTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLAAARNPFNCVCPLSWFG  
PWVRESHVTLASPEETRCHFPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALS  
SSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPE  
GFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGLQRYLQGSSVQLRSLRL  
TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPA  
VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPAG  
PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

**FIGURE 33**

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT  
CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**GAGACTCATCAG  
AAACATTTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG  
AAGAAAGGGAACTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG  
ACCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTCAGA  
TTTTCATTTCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG  
ATCTCAAAACCTTTGAATTTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG  
AAGAGTGTAACTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAATGACTT  
TGACACCATGCCTATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGA  
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC  
TTCTTAGGATTCAGAACTCTTCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC  
AAACTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGCCTGATGGAATCA  
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA  
ATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA  
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTTGGCATACATCAGTGGAACTACT  
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC  
TCAAATACTGTAATGAGAACTATAAAATGGAGCATGTACATTTCAGAGTGTTTTACATTCA  
ACAGGATAAAATCTATTTGCTTTTGGACCAAAATGGACATAGAAAACCTGACAAATCAAAATG  
CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATTCGAATATTTAAATTTTGGC  
AATAATATCTTAAACAGACGAGTTGTTTAAAAGAACTATCCAATGCCTCACTTGAAAACCTCT  
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC  
CCTTGGAACTTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTGCTCA  
TGGCCAGAACTGTGGTCAATATGAATCTGTCTACATAAATAAATTGTCTGATTCTGTCTTCAG  
GTGCTTGCCCCAAAAGTATTCAAATACTTGACCTAAATAAACCAAATCCAACTGTACCTA  
AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAACTGAT  
CTCCCTGGATGCAGTCATTTTCTAGTAGACTTTTCTGAACTTTGAAATGAACTTTCATTCT  
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAA  
ATCCATTCCGGTGTACCTGTGAATTAATAAATTTTCTTTCAGCTTGAAACATATTCAGAGGTC  
ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAAGTAG  
GTTAAAAGACGTTTCTCTCCACGAATTATCTTGCAACACAGCTCTGTTGATTGTCACCATTG  
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC  
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA  
AGAACAACCTCAAGAGAAATGTCCGATTCACAGCATTTTATTTTATACAGTGAACATGATTCTC  
TGTGGGTGAAGAATGAATTGATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTGTC  
CTTTATGAAAGCTACTTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA  
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGAGTGGTGCCATT  
ATGAATTTCTACTTTTGCCACCACAATCTCTTCCATGAAAATTCTGATCATATAATTCTTATC  
TTACTGGAACCCATTCCATTCTATTGCAATTTCCACCAGGTATCATAACTGAAAAGCTCTCCT  
GGAAAAAAAAGCATACTTGAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA  
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA  
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT  
**ATAA**AATCCCACAGTCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA  
CAACCTTTTATGATGGCAATTTGACAATATTTATTTAAAATAAAAAATGGTTATTTCCCTTCATA  
TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG  
GCTTATGGAAAAGGTGTTTATCCCAGGATTGTTTATAATCATGAAAATGTGGCCAGGTGC  
AGTGGCTCACTCTTGTAATCCCAGCACTATGGGAGGCCAAGGTGGGTGACCCACGAGGTCAA  
GAGATGGAGACCATCCTGGCCCAACATGGTGAAACCTGTCTCTACTAAAAATACAAAATTA  
GCTGGGCGTGATGGTGCACGCCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG  
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT  
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAATGGAAAACATCC  
TCATGGCCACAAAATAAGGTCTAATTCATAAATTATAGTACATTAATGTAATATAATATTA  
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTTGGTATGGAAAAACATATTAAATAT  
GTTATAAACTATTAGGTTGGTGCAAAACATAATTGTGGTTTTTGGCATTGAAATGCAATTGAA  
ATAAAAGTGTAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTTGGA  
TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTGTTTAGAATGA  
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT





**FIGURE 36**

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPPLP  
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS  
VRVAREDVRESCPPLGLETCLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA  
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIF  
EGNSEKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT  
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF  
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG  
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA  
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSAST  
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES  
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRL  
NPGEYVVTAKEGFTASTKNCMVGYDMGATRCDFTLSKTNMARIREIMEKFGKQPVSLPARR  
LKLGRGRRRQRG



**FIGURE 37**

CTAAGAGGACAAGATGAGGCCCGGCCTCTCATTTCTCCTAGCCCTTCTGTTCTTCCCTTGCCCAAGCTGCAGGGG  
 ATTTGGGGGATGTGGGACCTCCAATTTCCAGCCCCGGCTTCAGCTCTTTCCAGGTGTTGACTCCAGCTCCAGC  
 TTCAGCTCCAGCTCCAGGTGCGGCTCCAGCTCCAGCCGCAGCTTAGGCAGCGGAGGTTCTGTGTCCAGTTGTT  
 TTCCAATTTACCCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC  
 CCGTGGACAGAGTGGAACGCTTGGAATTCACAGCTCATGTTCTTTCTCAGAAAGTTTGAGAAAGAACTTTCTAAA  
 GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTTAAACCTAACTGTCCGAATTGACATCAT  
 GGAGAAGGATAACCATTTCTTACACTGAACCTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC  
 TGGTCATACAGCTGAAGGAGAGTTTTTGGTGGAAGCTCAGAAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT  
 ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTCGCCGAGAAATCGTGGC  
 TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTCGTCCACCCCTCTCCCACTC  
 CAGGGAGCTGTGGTTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTTCAGCTCAACTGGAGAGGGTTT  
 TCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCCAGCATCCAAACAAAGGACTGTATTGGGTGGCGCC  
 ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATTGTATATAA  
 ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC  
 ATGTACAACACCGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACCTCTCCCTAA  
 TGCTGCCTATAATAACCGCTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG  
 GATTGTGGGTATTTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT  
 CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGTTCT  
 GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAGG  
 GCAAACCTAGACATTGTAATGCATAAGATGCAGGAAAAAGTCAGAGCATTAACTATAACCCCTTTGACCAGAAA  
 CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAGTAAAGCTGTTTA  
 GGAGTTAGGGTGAAAGAGAAAATGTTTGTGAAAAAATAGTCTTCTCCACTTACTTAGATATCTGCAGGGGTGT  
 CTAAAAGTGTGTTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT  
 TGATTTGGTGAGTTCTCTTGGGAATCATCTGCCTCTTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGA  
 TTGTGAGAGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA  
 GGAATTAAGGAACCTAAACCTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC  
 CTCATCCATGTAGCACCACCTAATTCTTCCATGCCTGGAAGAAACCTGGGGACTTAGTTAGGTAGATTAATATCT  
 GGAGCTCCTCGAGGGACCAAATCTCCAACCTTTTTTTTCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG  
 CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTT  
 ATGCATTAAATTTGTACATGGCAAATAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTCTCTC  
 ATTTGTCCACCTTACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCCTTCCAAAGGCAGCTCAGAAGATTAG  
 AACCAGACTTACTAACCAATTCCACCCCCACCAACCCCTTCTACTGCCTACTTTAAAAAAATTAATAGTTTT  
 CTATGGAAGTATCTAAGATTAGAAAAATTAATTTCTTTAATTTCAATTATGGACTTTTATTTACATGACTCTA  
 AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA  
 TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTTAAGCCTGGAACCTGTAAGAAAAT  
 GAAAATTTAATTTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT  
 TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT  
 TCCTTTGATGTTCAAGTCCTAGTCTATAGGATTGGCAGTTTAAATGCTTTACTCCCCCTTTTAAATAAATGAT  
 TAAATGTGCTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 38**

MRPGLSFLLALLFFLGQAAGDLGDVGPPIPSPGFSSFPGVDSSSSFSSSSSRSGSSSSRSLGS  
GGSVSQQLFSNFTGSDDRGTCQCSVSLPDTTFPVDRVERLEFTAHVLSQKFEKELSKVREYV  
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ  
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTKLKECEASKDQNTPVVHPPPTPGSCGH  
GGVVNISKPSVVQLNWRGFSYLYGAWGRDYSPQHFNKGLYWVAPLNTDGRLLLEYRLYNTLD  
DLLLYINARELRITYGQSGTAVYNNMYVNMVNTGNIARVNLTNTIAVTQTLPNAAAYNNR  
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDTTLQVLNTWYTKQYKPSASNAF  
MVCGLVLYATRMTNTRTEEIFYYYDTNTGKEGKLDIVMHKMQEKVQSINYNPFDQKLYVYNDG  
YLLNYDLSVLQKPQ

1000 999 998 997 996 995 994 993 992 991 990 989 988 987 986 985 984 983 982 981 980 979 978 977 976 975 974 973 972 971 970 969 968 967 966 965 964 963 962 961 960 959 958 957 956 955 954 953 952 951 950 949 948 947 946 945 944 943 942 941 940 939 938 937 936 935 934 933 932 931 930 929 928 927 926 925 924 923 922 921 920 919 918 917 916 915 914 913 912 911 910 909 908 907 906 905 904 903 902 901 900 899 898 897 896 895 894 893 892 891 890 889 888 887 886 885 884 883 882 881 880 879 878 877 876 875 874 873 872 871 870 869 868 867 866 865 864 863 862 861 860 859 858 857 856 855 854 853 852 851 850 849 848 847 846 845 844 843 842 841 840 839 838 837 836 835 834 833 832 831 830 829 828 827 826 825 824 823 822 821 820 819 818 817 816 815 814 813 812 811 810 809 808 807 806 805 804 803 802 801 800 799 798 797 796 795 794 793 792 791 790 789 788 787 786 785 784 783 782 781 780 779 778 777 776 775 774 773 772 771 770 769 768 767 766 765 764 763 762 761 760 759 758 757 756 755 754 753 752 751 750 749 748 747 746 745 744 743 742 741 740 739 738 737 736 735 734 733 732 731 730 729 728 727 726 725 724 723 722 721 720 719 718 717 716 715 714 713 712 711 710 709 708 707 706 705 704 703 702 701 700 699 698 697 696 695 694 693 692 691 690 689 688 687 686 685 684 683 682 681 680 679 678 677 676 675 674 673 672 671 670 669 668 667 666 665 664 663 662 661 660 659 658 657 656 655 654 653 652 651 650 649 648 647 646 645 644 643 642 641 640 639 638 637 636 635 634 633 632 631 630 629 628 627 626 625 624 623 622 621 620 619 618 617 616 615 614 613 612 611 610 609 608 607 606 605 604 603 602 601 600 599 598 597 596 595 594 593 592 591 590 589 588 587 586 585 584 583 582 581 580 579 578 577 576 575 574 573 572 571 570 569 568 567 566 565 564 563 562 561 560 559 558 557 556 555 554 553 552 551 550 549 548 547 546 545 544 543 542 541 540 539 538 537 536 535 534 533 532 531 530 529 528 527 526 525 524 523 522 521 520 519 518 517 516 515 514 513 512 511 510 509 508 507 506 505 504 503 502 501 500 499 498 497 496 495 494 493 492 491 490 489 488 487 486 485 484 483 482 481 480 479 478 477 476 475 474 473 472 471 470 469 468 467 466 465 464 463 462 461 460 459 458 457 456 455 454 453 452 451 450 449 448 447 446 445 444 443 442 441 440 439 438 437 436 435 434 433 432 431 430 429 428 427 426 425 424 423 422 421 420 419 418 417 416 415 414 413 412 411 410 409 408 407 406 405 404 403 402 401 400 399 398 397 396 395 394 393 392 391 390 389 388 387 386 385 384 383 382 381 380 379 378 377 376 375 374 373 372 371 370 369 368 367 366 365 364 363 362 361 360 359 358 357 356 355 354 353 352 351 350 349 348 347 346 345 344 343 342 341 340 339 338 337 336 335 334 333 332 331 330 329 328 327 326 325 324 323 322 321 320 319 318 317 316 315 314 313 312 311 310 309 308 307 306 305 304 303 302 301 300 299 298 297 296 295 294 293 292 291 290 289 288 287 286 285 284 283 282 281 280 279 278 277 276 275 274 273 272 271 270 269 268 267 266 265 264 263 262 261 260 259 258 257 256 255 254 253 252 251 250 249 248 247 246 245 244 243 242 241 240 239 238 237 236 235 234 233 232 231 230 229 228 227 226 225 224 223 222 221 220 219 218 217 216 215 214 213 212 211 210 209 208 207 206 205 204 203 202 201 200 199 198 197 196 195 194 193 192 191 190 189 188 187 186 185 184 183 182 181 180 179 178 177 176 175 174 173 172 171 170 169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154 153 152 151 150 149 148 147 146 145 144 143 142 141 140 139 138 137 136 135 134 133 132 131 130 129 128 127 126 125 124 123 122 121 120 119 118 117 116 115 114 113 112 111 110 109 108 107 106 105 104 103 102 101 100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0







**FIGURE 42**

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT  
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC  
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAACAANTCCACTGTAAC TAGATTGATCTA  
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG  
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT  
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT  
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAAATGGAT  
TTTGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC



**FIGURE 44**

AAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGTGAACGTTNTGGAGATGGGGAGC  
GTCCTTGGGGTTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCCCCGTGTT  
TGCTATGCCGATGCTGTCCTAGTGGAAACAACCTCCACTGTAAC TAGATTGATCTATGCACTT  
TTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAAC  
GAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATTTTGGTTG  
GCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTCTTA  
CTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTT  
CTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC



**FIGURE 45**

GCTGTCCTTAGTGGAAACAANTCCAACCTTGTAACCTGGATTGATCTATGCACTTTTTTCCTTG  
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA  
AGATTCCTGGATTTTTGTGAGAATGAGAAAGGTGTTGTCCCCTTGTAACATTTTTGGTTGGC  
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTCTTTACT  
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT  
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACTTTT  
ACAACTGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTTGTTTCATCCTCATACAAGTAGT  
CTTACTTATTGATTTTGCACATTCATGGAATGAATCGTGGGTGAAAAAATGGAAGAAGGGA  
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA  
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC  
GTTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

bioRxiv preprint doi: <https://doi.org/10.1101/2021.03.11.434111>; this version posted March 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 46**

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCTGCGATTGAGCTGCGGGTCGCGGCCGGCGCCGGCCTCTCCAAT  
 GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTTCGGCAAAGGCAGTCGAGTGTTGCAGACCGGGGCGAG  
 TCCTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCGCGGGGCTGTGCG  
 ACTCCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTAC  
 GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC  
 CCCTGGTTTGTGTGTTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCC  
 TGGGCGAATCCCACATCTGTTTCACTCTCCGCCGAGGGCGAGCAGGAGCGGAGAGTGTGTGCAATCTGCGAGTG  
 AAGAGGGACGAGGGAAAAGAAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGAT  
 CAGCAAAAAAGAAGATGGGGCCCCCGAGCCTCGTGCTGTGCTTGCTGTGTCGCAACTGTGTTCTCCCTGCTGGG  
 TGGAAGCTCGGCCCTTCTGTGCGACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCA  
 ACATCATCTCTGGTGTGACGGACGACCAGGATGTGGAGCTGGGTTCATGCGAGGTGATGAACAAGACCCGGCGC  
 ATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCTCACGCTCCTC  
 CATCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC  
 AGGCACAGCAGGAGCGCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG  
 TATCTTAATGAATACAACGGCTCCTACGTGCCACCGCTGGAAGGAGTGGGTGCGACTCCTTAAAACTCCCG  
 CTTTTATACTACAGCTGTGTGCGGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA  
 CAGACCTCATACCAATGACAGCGTGAGCTTCTTCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCTC  
 ATGGTCATCAGCCATGCAGCCCCCAGGCCCTGAGGATTACGCCCCACAATATTACGCCTCTTCCAAACGC  
 ATCTCAGCACATCAGCCGAGCTACAACACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC  
 CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGACCCATGTGCGGTGGAC  
 GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC  
 CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG  
 TCCCGTTCTACGTGAGGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG  
 GCCCCACCATCCTGGACATTCGAGGCCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCT  
 GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG  
 AGAGAGGCAAGCTGTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTAC  
 CAGCGTGTGAAGGACCTGTGTGAGCGTGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG  
 TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC  
 TCTCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCGTGCACCTGTGACAGCGGGGACTACAAGCTC  
 AGCCTGGCCGACGCCGGAACCACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCACTCGCTCCATCCG  
 CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA  
 CCAAGCGCACTGGCCAGGGGGCCCCGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC  
 CTTCCCGACTACTCAGCCGCCAACCCCATTTAAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA  
 GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCCTGGAAGACCACAAGCTGCACATCGACCACGAGATTGAAA  
 CCCTGCAGAACAAAATTAAGAACCTGAGGGAAGTCCGAGGTACCTGAAGAAAAAGCGGCCAGAAGAATGTGAC  
 TGTACAAAATCAGCTACCACACCCAGCACAAAGGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTTCAG  
 GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGC  
 TCAAGCGCCTGCAGAACACGACAGCTGCAGCATGCCAGGCCACAGTGTTCACCCACGACAACCAGCACTGG  
 CAGACGGCGCCTTTCTGGACACTGGGGCCTTCTGTGCTGCACCAGCGCCAACAATAACACGTACTGGTGCAT  
 GAGGACCATCAATGAGACTCACAATTTCTCTCTGTGAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCA  
 ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCTCAACCAGCTACACGTACAG  
 CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGAACCCCCGAGCTCGAAACATGGACCTGGATGGAGG  
 AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG  
 GACAACTGTGGGAAGGCTGGGAAGGTAAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA  
 CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC  
 AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC  
 ATTTTTGCCCTGCTTTTGGCTTTGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTC  
 ACCACTAACCCTCCCCAGAAAGCTCACAAAGGAAAACGGAGAGAGCGAGCAGAGAGATTTCTTGGAAATTTTC  
 TCCCAAGGGCGAAAGTCATTGGAATTTTTTAAATCATAGGGGAAAAGCAGTCTGTCTAAATCCTCTTATTCTT  
 TTGGTTTGTACAAAAGAAGGAACATAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG  
 TTTGACAAATGAGTCAGTAGCACAAAAGAGATGACATTACCTAGCACTATAAACCTGGTTGCCTCTGAAGAAA  
 CTGCCCTTCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACCTTTAGGGGAACCTAATAAGAAAT  
 CCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAGAAAA

**FIGURE 47**

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ  
VMNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHE  
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD  
YSKDYLTDLITNDSVSFFRTSKKMPHRPVL MVISHAAPHGPEDSAPQYSRLFPNASQHITP  
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT  
YIVYTADHGYHIGQFGLVKGKSMPEFDIRVPFYVRGPVNEAGCLNPHIVLNIDLAPTILDI  
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN  
FLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY  
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKKYKASYVRSRSIRSV AIEVDGRVYHVGLGDAAQ  
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS  
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL  
HPFRKGLQEKDKVWLLREQRKKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG  
PFCACTSANNNTYWCMRTINETHNFLCFEATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL  
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA  
GCTGCTGGGCAGAGAGGGGACTGTCCGGCTCCCAG**ATG**CTGGGCCTCCTGGGGAGCACAGCCC  
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCTGCTGCTGCTGCTGCTGCTGGCCACC  
TGCCTTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGGAAA  
CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGCGGGGCCACCTGGGAATCTTTCACC  
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC  
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA  
CGCTCGC**TGA**GGCTGCTGTGCGCCGTGCCTGTGGACAGCAGCTGCCCTGCCCTCCCATCTG  
TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG  
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC  
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG  
GTGCTCTTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAGTGAGCCCCAGAAATG  
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC  
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT  
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTTCATCGGTGCCTTAGTCCAAGAAAAT  
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA



**FIGURE 50**

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCC**ATG**TCGGACCTGCTA  
CTACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG  
GTACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCA  
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC  
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC  
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC  
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC  
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTCTGTCCATCTGGCTGGCTACCCGCCG  
TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG  
AGATCTACCAGGAAGACCAGATCCATTTTATGTGCCCACTGGCACGGCAGGGAGACTTCTAT  
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA  
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC  
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT  
GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA  
GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC  
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAG**TAA**CCC  
ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT  
CCAGCCCTCTTCCTCCTTCTCTGGGGGAGGAGGGGTTCTGAGGGACCTGACTTCCCCTGC  
TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA  
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTCAGACTC  
ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAA  
AAAAA

**FIGURE 51**

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGlyGETGR  
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS  
FPAPSHVVTATFPYTTILSIWLATRVRHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPAR  
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSATLSPGAS  
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK  
GKE

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGlyGETGR  
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS  
FPAPSHVVTATFPYTTILSIWLATRVRHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPAR  
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSATLSPGAS  
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK  
GKE

**FIGURE 52**

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCT  
GCCCCGCGCCCAGTC**ATG**ACCCTGCGCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT  
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA  
CCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGA  
GACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT  
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA  
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTAT  
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT  
GATTGCACTAATCCGAGCCAACACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAG  
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAAT  
AGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAA**TA**  
**A**TAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA



**FIGURE 53**

MTLRPSLLPLHLLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFCDTLHI  
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF  
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS  
KKKLKEEKRNKSKKK

MTLRPSLLPLHLLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFCDTLHI  
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF  
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS  
KKKLKEEKRNKSKKK



**FIGURE 55**

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC  
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG  
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGA ACTTGGCCAAAAGCAGGTGA  
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT  
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT  
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG  
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC  
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAA  
CAAGAGCAAAAAGAAATAATAATAATAAATTTTAAAAAACTTAAAA



57/330

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTHASIAAGVLFFTGWWIIIDA AVIYPTMKDFNHSYHACGVI  
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK  
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

Category	Sub-category	Value	Unit	Value	Unit	Value	Unit	Value	Unit
Energy	Electricity	1000	kWh	1000	kWh	1000	kWh	1000	kWh
	Gas	1000	kWh	1000	kWh	1000	kWh	1000	kWh
Water	Hot water	1000	kWh	1000	kWh	1000	kWh	1000	kWh
	Cold water	1000	kWh	1000	kWh	1000	kWh	1000	kWh
Waste	Recycling	1000	kWh	1000	kWh	1000	kWh	1000	kWh
	Landfill	1000	kWh	1000	kWh	1000	kWh	1000	kWh



**FIGURE 59**

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTTCAGAATGCATGAC  
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG  
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC  
ATGCCTGTGGTGTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA  
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT  
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTTGGAGGTT  
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC  
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG





**FIGURE 61**

TAGAGGGCCTTGAGATGCTCAGAAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC  
ATTGNTGNTGGTGTANTATTTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT  
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC  
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG  
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTGGATCTN  
TGATTGCATTTATGTGGATTNTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC  
CCTGT

**FIGURE 62**

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCCGGCGCGG  
CGNGACACCGGGTTCCGGGAACCATTGACACGACGGGGTGGACTGACCTGAAAAAAATGTTTG  
GATTTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT  
GCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT  
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCA  
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA  
GGTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGG  
ATNTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG  
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTGGAGGGCTG

**FIGURE 63**

CGACGCCGGCGTG**ATG**TGGCTTCCGCTGGTGCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCC  
TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC  
AAACGGCCCCCAGCGCCCCTGGTAAGTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC  
TTTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG  
GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGCTGGAACAACAT  
ACCAAGGCAGGGGGCTGCTGTGCATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT  
CCATTACATTGGGCGTATGGAAGAGGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG  
AAGGGCAGCTGGACTGGGCTCCCCCTGTCTCTCCTTTTGACATCATGGTACTGGAAGGGCCC  
AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTTCAGGGCCTCAAGGA  
GAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA  
GTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTTGCCCGTGGTTCAGCTCCTCGAC  
AGGTGTGGGCTGCTGACTCGTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGA  
GGTCTGCAGCAGCTGGGGCCCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA  
CTTACGGTGTACCCCCAACCACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC  
ATGAAAGGAGGCTTTTATCCCCGAGGGGGTTCCAGTGAAATTGCCTTCCACACCATCCCTGT  
GATTTCAGCGGGCTGGGGGGCGCTGTCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT  
CAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC  
CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAAAGC  
CCGCTGCCTGCCAGGTGTGAAGCAGCAACTGGGGACGGTGGCGCCCGGCTTAGGCATGACCT  
CTGTTTTTCATCTGCCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT  
GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCTCCATGCCAGGGGAAGA  
GGCTGCGGAACACATCCCTCTTCTCTTCTCGCTTTCCCATCAGCCAAAGATCCGACCTGGG  
AGGACCGATTCCCAGGCCGGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTTT  
GAGGAGTGGCAGGCGGAGCTGAAGGGAAAGCGGGGCAGTGAAGTATGAGACCTTCAAAAACCTC  
CTTTGTGGAAGCCTCTATGTGAGTGGTCTGAAACTGTTCCACAGCTGGAGGGGAAGGTGG  
AGAGTGTGACTGCAGGATCCCCACTCACCAACCAGTTCTATCTGGCTGCTCCCCGAGGTGCC  
TGCTACGGGGCTGACCATGACCTGGGCCGCCTGCACCCTTGTGTGATGGCCTCCTTGAGGGC  
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACTGTGGACTGGTCG  
GGGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACCTTGTACTCAGAC  
CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAT**TAG**TTCCATCAGGGAGG  
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG  
CATTAGTTCTTGCACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG  
TTTAAATCACAAATCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA  
TCTTTACGCCTTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATAGCTTGATG  
TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCTCCTAACTCAGTGATCAAAGCGA  
ATATTCCATCTGTGGATAGAACCCTGGCAGTGTGTCAGCTCAACCTGGTGGGTTCAGTTC  
TGTCTGAGGCTTCTGCTCTCATTCAATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG  
GAAAAGGGAGACTAATGAGGCTTAACCTAAAACCTGGGCGTGGTTTTGGTTGCCATTCCATA  
GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTTCAGTGGCTCTTCAGGGGACAGGAAT  
GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA  
GGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCGGGTCTGTCTTCCTTATCA  
TCGGGGTGGCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGC  
CTTATCCACCAATAACACAGGGAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCA  
TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG  
CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGGAAGTCACATCAGAAAAGGGA  
AAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTTCAGTTAATTAGCACATGTGAGGG  
TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTG  
GTATCAGACATACGAAAGGTCTCTTTGTAGTTCTTAATGTAACATTAATAAATTTATTG  
ATTCCATTGCTTTAAAAA

**FIGURE 64**

MWLPLVLLLLAVLLLLAVLCKVYLGLEFSGSSPNPFSEDKRPPAPLVTDKEARKKVLKQAFSAN  
QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG  
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ  
EEAIIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ  
LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHMKGGFYPRGGSSEIAFHTIPVIQRA  
GGAVLTkATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP  
GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTMDQAMERYVSMPREEEAAEH  
IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAELKGKRGSDYETFKNSEVEA  
SMSVVLKLFPPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI  
PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

**FIGURE 65**

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA  
 GGGGTTGGCACC GGCCCCGAGAGGAGG**ATG**CGGGTCCGGATAGGGCTGACGCTGCTGCTGTG  
 TGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT  
 TAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTAAAGGACCATACTACTGCAGGCAGA  
 GTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGAATTAGAATCCTCTATTCAAGA  
 AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAGAAGATATCAGCTTTCTAG  
 AGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTACGGAAACCAGCTTTG  
 ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCTAGATAA  
 GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT  
 ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCTAAGAGA  
 CGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAGCAA  
 TAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAGGCAGCAAGCATGAACCATA  
 CCAAAGCCCTGGAGAGAGTGTATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC  
 CAGGCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC  
 TCTTGGCTTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAGTCAGGCAAAGGCTCTTGTAT  
 ATTATACATTTGGAGCTCTTGGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTT  
**TAG**TGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTT  
 TCAGCTTTCATGATCCAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACTTCCTTCAA  
 ATTCCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA  
 ATTTTTCTTTAAAATGATTAGTTTGGCTGATTGCCCCATAAAAGAGAGATCTGATAAATGGC  
 TCTTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTTACATTAGATTATCATAA  
 TTTTAAAAATTTTCTTTAGTTTTTCAAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT  
 GAAATATTATACAATATTTTGCACAATGCCCTAAGAATTGTTAAATTCATGGAGTTATTT  
 GTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATTGGCTGTCTTC  
 CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC  
 CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACATTTTTTAATAAA  
 ATTATGTCTAAGATTAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSKTTLTSDSESVKDHTTAGRVVAGQIFLD  
SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG  
EPCHFPPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMQEAEMM  
YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEEK  
LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAAACTTCTACTTGCCT  
GCCCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT  
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG  
CTCAACTCTCCTGCACGCTCAGCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGG  
TACCAGCAGCGGGCAGGCAGTGCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA  
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCAACAATGCCT  
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC  
TACGGCTTTAGTCCCT**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTCT  
GCCCCTGACCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAATGGG  
TTAATAATATTCAACATGTCAACAAC

## 113

MACRCLSFLMGTFLSVSQTVLAQLDALLVFPQGVAQLSCTLSPQHVTIRDYGVSWYQQRAG  
SAPRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVPEDDADYYCSVGYGFSF









**FIGURE 72**

MVGTKAWVFSFLVLEVTSVLGRQTM LTQSVRRVQPGKKNPSIFAKPADTLES PG EWTTWFNI  
DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ  
RPGQNC SNYTVRFLCPPGSLRRDTERI WSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS  
EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGA AIYLLTK  
TPKLLTQTDS DGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTS LKAATIKAEFVRAETPY  
MVMNPETKARRAGQSVSLCKKATGKPRPD KYFWYHNDTLLDPSLYKHESKLVLRLK LQQHQAG  
EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQ NATNSFY YDVGRCPV  
KTCAGQQDNGIRCRDAVQNC CGISKTEEREIQCSGYTLPTKVAKECSCQRCTETR SI VRGRV  
SAADNGEPMRFGHVYMGNSRVSM TGYKGTFTLHVPQDTERLVLTFVDRLQKFVNTTKVLPFN  
KKGS AVFHEIKMLRRKEPIT LEAMETNIIPLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKV  
KASVTFLDPRNISTATAAQTD LNFINDEGDTFPLRTYGMFSVD FRDEVTSEPLNAGKV K VHL  
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL  
DVPESRRRCFVKVRAYRSE RFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA  
CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPR  
VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN  
EDDPMSWTE DYLAWWPKPMEFRACYIKVKIVGPLEVNVR SRNMGGTHRRTVGKLYGIRDVRS  
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVIPQGSCRRASVNPM LHEYLVNHLPLAV  
NNDTSEYTM LAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT  
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA  
QQPLIN







76/330

## **FIGURE 76**

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF  
SSFQDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQVKW  
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFERMEPVTA  
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF  
SSFQDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQVKW  
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFERMEPVTA  
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS







**FIGURE 79**

MEAPDYEVL SVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNK  
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL  
IFLMPFAYFFTESEGFAGSRKGVLG RVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL  
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE  
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLEKRRKASAWQRNLGYPLAMLCLLVLT  
GLSVLIVAIHILELLIDEAAMPGRMQGTS LGQVSFSKLGSEFGAVIQVVLI FYLMVSSVVG FY  
SSPLFRSLRPRWHD TMTQII GNCVCLLVSSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY  
IVFLYNAAFAGLTTLCLVKTF TAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

MEAPDYEVL SVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNK  
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL  
IFLMPFAYFFTESEGFAGSRKGVLG RVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL  
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE  
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLEKRRKASAWQRNLGYPLAMLCLLVLT  
GLSVLIVAIHILELLIDEAAMPGRMQGTS LGQVSFSKLGSEFGAVIQVVLI FYLMVSSVVG FY  
SSPLFRSLRPRWHD TMTQII GNCVCLLVSSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY  
IVFLYNAAFAGLTTLCLVKTF TAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

**FIGURE 80**

GGCTGCCGAGGGAAGGCCCTTGGGTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC  
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC  
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA  
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC  
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG



**FIGURE 82**

GATGTGCTCCTTGGAGCTGGTGTGCAGTGCCTGACTGTAAGATCAAGTCCAAACCTGTTTT  
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTCATGCTGCTGT  
GGGTGATATTACTGGTCCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCATT  
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA  
GGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA  
TACTAAGAGAAACCCCAGACAATATCCTTGAGGTTGAGGAATCTGGAGAGTACAGATGCCAG  
GCCCAGGGCTCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC  
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCTTAGGCCTCTC  
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC  
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA  
TGATAATGTCCTGGCATTCCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA  
AAA



**FIGURE 84**

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGT  
 GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG  
 GAGTGTGTGGAACAGGACCCGGGACAGAGGAACC**ATG**GCTCCGCAGAACCTGAGCACCTTTT  
 GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG  
 GGGGTGCCTCGAAGTGCTCTATAAAGGATATTA AAAAGGCCTATAGGAACTAGCCCTGCA  
 GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG  
 CTGCTTATGAGGTTCTGTGAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA  
 GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTCTTTGGGGATTT  
 TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTTCCAAGAGGAAGTGATA  
 TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT  
 AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTGCGCAAGAGAT  
 GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT  
 GCCCTAATGTCAAACCTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG  
 AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG  
 AGATTTACGGTTCGAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATT  
 TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT  
 CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT  
 ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA  
 TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAA  
 CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TG**AGAGTG  
 AATAAAATTGGACTTTGTTTAAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTT  
 TTGTGTGTGTTTTTGTTTTTATTTTCAATATGCAAGTTAGGCTTAATTTTTTTTATCTAATGA  
 TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTTCGGAAAAGAATGACC  
 AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT  
 TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA  
 GTTGTAGCAATTTCAATCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTG  
 TTATTTTAA



**FIGURE 85**

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ  
 AQEKFQDLGAAAYEVLSDEKRKQYDITYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGTPRQQ  
 DRNIPRGSDIIVDLEVTLLEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ  
 MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGM EY PFIGEGEPHVDGEPGDLRFRIKVVKH  
 PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD  
 NNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

**Important features:****Signal peptide:**

amino acids 1-22

**Cell attachment sequence.**

amino acids 254-257

**Nt-dnaJ domain signature.**

amino acids 67-87

**Homologous region to Nt-dnaJ domain proteins.**

amino acids 26-58

**N-glycosylation site.**

amino acids 5-9, 261-265

**Tyrosine kinase phosphorylation site.**

amino acids 253-260

**N-myristoylation site.**

amino acids 18-24, 31-37, 93-99, 215-221

**Amidation site.**

amino acids 164-168

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ  
 AQEKFQDLGAAAYEVLSDEKRKQYDITYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGTPRQQ  
 DRNIPRGSDIIVDLEVTLLEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ  
 MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGM EY PFIGEGEPHVDGEPGDLRFRIKVVKH  
 PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD  
 NNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

[illegible][illegible]

**FIGURE 87**

GGCACGAGGCGGCGGGGCGAGTCGCGGGATGCGCCCCGGGAGCCACAGCCTGAGGCCCTCAGGT  
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA  
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATCG**  
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTCTGGCTTCGTTTGCAGCC  
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC  
TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC  
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC  
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC  
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA  
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG  
TACCCTCCGTTGGACCCCCAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTCTAG  
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC  
AGTCTCTGTCTGGCTGCTGAGGAGCATTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG  
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTT**TAG**GCCT  
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT  
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG  
TAAAGCAGGAGATCCCCGTAGTTTATGCCTCTTTTGCAGTTGCAAACGTGGCTGGTGAGT  
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA  
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT  
AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG  
TGTGGCAATTCTGATCTGCATTTTCAGAAGAGGACAATCAATTGAACTAAGTAGGGGTTTC  
TTCTTTTGGCAAGACTTGTA CTCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT  
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAAGGTTTGGGTTTGAAGCTGAGGAACT  
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG  
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAA

Case	Age	Sex	Site	Pathologic	Survival	Follow-up
1	62	M	Rectum	Adenocarcinoma	10 years	10 years
2	58	F	Rectum	Adenocarcinoma	10 years	10 years
3	65	M	Rectum	Adenocarcinoma	10 years	10 years
4	60	F	Rectum	Adenocarcinoma	10 years	10 years
5	63	M	Rectum	Adenocarcinoma	10 years	10 years
6	61	F	Rectum	Adenocarcinoma	10 years	10 years
7	64	M	Rectum	Adenocarcinoma	10 years	10 years
8	62	F	Rectum	Adenocarcinoma	10 years	10 years
9	66	M	Rectum	Adenocarcinoma	10 years	10 years
10	60	F	Rectum	Adenocarcinoma	10 years	10 years
11	63	M	Rectum	Adenocarcinoma	10 years	10 years
12	61	F	Rectum	Adenocarcinoma	10 years	10 years
13	64	M	Rectum	Adenocarcinoma	10 years	10 years
14	62	F	Rectum	Adenocarcinoma	10 years	10 years
15	65	M	Rectum	Adenocarcinoma	10 years	10 years
16	60	F	Rectum	Adenocarcinoma	10 years	10 years
17	63	M	Rectum	Adenocarcinoma	10 years	10 years
18	61	F	Rectum	Adenocarcinoma	10 years	10 years
19	64	M	Rectum	Adenocarcinoma	10 years	10 years
20	62	F	Rectum	Adenocarcinoma	10 years	10 years

**FIGURE 88**

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL  
ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS  
VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLLSVSHLVLVTRNACHLTGGGLDWI  
DQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

**FIGURE 89**

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTCCGAGGTGCTTTCGCCGCTGTCC  
CCACCACTGCAGCCATGATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA  
TTTGGAGTGTTTTTCCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT  
TGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTCAGAT  
TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC  
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTCTCTTGTTTCAG  
GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT  
TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTATTAACAACA  
AGTGAATTTGAAGACTCATTTAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA  
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG  
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA  
ACTAAGAAGAAGTCAGCAAGCAAAGTCTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA  
ACTCTTGAAGGCTATTTGTGTTGTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA  
CTGTGGTGCCTGTTTCTTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT  
TTTTAGAAGTGCTCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA  
TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAACCAAGGAAACCCCAATTTTG  
ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

90/330

## **FIGURE 90**

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK  
HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFI RRV PVLGSLNLPGI  
RSFVDKVGESNNMV

### **Important features:**

#### **Transmembrane domains:**

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

#### **N-myristoylation sites.**

amino acids 11-16, 51-56 and 116-121

#### **Aminoacyl-transfer RNA synthetases class-II protein.**

amino acids 49-59

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200

**FIGURE 91**

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC  
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA  
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTTCCTGTTCTTTGGA  
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT  
GGCTTTTGTAATTGGTTTAGAAAGAACATTCAGATTCTTCTTCCAAAACATAAAATGAAAG  
CTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG  
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

1000 900 800 700 600 500 400 300 200 100 0  
G A A G A C G T G G C G G C T C T C G C C T G G G C T G T T T C C C G G C T T C A T T T C T C C C G A C T C A G C T T C C C  
A C C N T G G G C T T T C C G A G G T G C T T T C G C C G C T G T C C C C A C C A C T G C A G C C A T G A T C T C C T T A A  
C G G A C A C G C A G A A A A T T G G A A T G G G A T T A A C C G G A T T T G G A G T G T T T T C C T G T T C T T T G G A  
A T G A T T C T C T T T T T G A C A A A G C A C T A C T G G C T A T T G G A A A T G T T T T A T T T G T A G C C G G C T T  
G G C T T T T G T A A T T G G T T T A G A A A G A A C A T T C A G A T T C T T C T T C C A A A A C A T A A A A T G A A A G  
C T A C A G G T T T T T T C T G G G T G G T G T A T T T G T A G T C C T T A T T G G T T G G C C T T T G A T A G G C A T G  
A T C T T C G A A A T T T A T G G A T T T T T C T C T T G T T C

**FIGURE 92**

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA  
GGCTGCCAGGAAGGAGACGCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGA  
CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC  
TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC  
ATCAACACCATTTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAAACAAGCAGCTCTTCCGGAA  
GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT  
CGGGCACGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAAT  
GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA  
ACGCTTTGGGCTGTTAGGGGGGCTCCAAGGTCTGGCCAAGAAAGAGCTGGCCTATGTCCCAA  
TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTTCGCGCAAGTGGGAGCAGGAT  
CGCAAGACGGTTGCCACCAGTTTGAGCAGCCTCCGGGACTACCCCGAGAAGTATTTTTTCTCT  
GATTCAGTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC  
GGGCCAAGGGGCTGCCTCGCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC  
ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTAGAAA  
TAATGAAAATCCAACACTGCTGGGAGTCTTAAACGGAAAGAAATACCATGCAGATTTGTATG  
TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC  
AAGCTCTACCAGGAGAAGGATGCCTTTTCAAGGAGGAGTACTACAGGACGGGCACCTTCCCAGA  
GACGCCCATGGTGCCCCCGGCGGCCCTGGACCCTCGTGAAGTGGCTGTTTTGGGCCCTCGC  
TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG  
CTGGCCAGCTTCATCCTCGTCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGT  
GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGACT**T**  
**G**ACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAAGTGGTGGCCTCTGCATATCCT  
CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA  
CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT  
TTTTCCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGC  
TGTGTGGTGAGTGTGAACCTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG  
GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCTTTTCATCCTTTGGTGCTGAGTTTTCTGT  
AACCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC  
CAAGAAAAAAAATTAAAGTGCTTTTCTGGGTCAAAAAAAAAA



[illegible][illegible]

**FIGURE 94**

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTGCGGCGGTGCTCTCGGGCTTTGTGCTCG  
 GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA  
 GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA  
 TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG  
 GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT  
 TGGTACAAATTCCGTCGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA  
 AACTTGCAGGAGCATTTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA  
 TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTT  
 TTTCACAGGGTACCTTTAGTGTTGCCAATCTGGGCATGTCTGAACAACCTGGGTTATAAAAC  
 TGTATCAGGTTCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT  
 TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA  
 CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT  
 AGTAAAGGATGTAAACAGATTAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG  
 CAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA  
 CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTTCATGTGTTATGTCTTTAAAAAATAGACA  
 TGTTTCTAAAAGTAGCTGTAACACCAACCACATCTCGATGTAGTAGACAATCTGACCTTAA  
 TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT  
 AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA  
 CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAATGAGCAGCC  
 CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTTGGTGAATATTCACGGTCTCCTACA  
 TTT**TGA**TCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT  
 TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC  
 CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC  
 ATCAGATGCTTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCT  
 TACACAGACACATTCTTTAGAAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA  
 TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC  
 TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT  
 TTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTTAAG  
 TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA  
 ATAAATTTGCAAAACATCATCTAAAATTTAAAAAAAAAAAAAAAAAAAAA

95/330

**FIGURE 95**

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIDIQ  
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH  
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC  
MSTGFSRAVQTHSSKFFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN  
RLKREIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS  
CNYNHHLDDVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLDLDQDKRSKA  
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIDIQ  
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH  
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC  
MSTGFSRAVQTHSSKFFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN  
RLKREIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS  
CNYNHHLDDVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLDLDQDKRSKA  
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF



**FIGURE 97**

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ  
GSTCCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF  
VKTYGHLYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEY  
LECVSKYTEQLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL  
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIES  
VMDPIDVKISDAIMNMQDNSVQVSQKV FQCGGPPKPLPAGRISRSISESAFSARFRPHHPEE  
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLEF  
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE  
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

**FIGURE 98**

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT  
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA  
GCAACTTACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGC  
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCAC  
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA  
AGATTGGTTTCTGAGAGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGC  
AGTGCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGA  
AAGCCAAACAAGCATTCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT  
TGCTCTGCCTTTGTAGGAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA  
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCCACCTCACTCTCCCCTGTACCCACC  
CCTAAATCATTCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTGTTGCTCTC  
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT  
AATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCTAGCTAGTGTCAATTAACCTTAAATGC  
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTTAAATGTCAAAAAAAAAAAAAAAAAA

[illegible][illegible]





101/330

**FIGURE 101**

MAVLVLRLTVVLGLLVFLTCYADDKPKDPDDKPDGKDPKPDFPKFLSLLGTEIENAVE  
FILRSMRSRSTGFMEFDDNEGKHSSK

MAVLVLRLTVVLGLLVFLTCYADDKPKDPDDKPDGKDPKPDFPKFLSLLGTEIENAVE  
FILRSMRSRSTGFMEFDDNEGKHSSK

**FIGURE 102**

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT  
CAGAGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTAC  
CCTGCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCC  
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG  
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG  
CTGCGGAACCGGAGCCAACCTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC  
CAAATCCCCACTTTGAGAAGTTCTTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT  
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT  
GGTGGTCTGCACTCTGGTGTCTGTGCTCTGTGCAGAGCCCCAAGGAAGGTCCTGCAGGAGGTCC  
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA  
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG  
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCAGTTCTCCGAAATCCAAATGG  
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC  
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA  
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTAGCAGAATGAGAGAAGACATT  
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC  
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC  
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC  
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT  
CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT  
CTCCCCACTACCACCTTCTTCCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT  
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA  
AACCACG

103/330

**FIGURE 103**

MDILVPLLQLLVLLLTPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKREIFSQI  
KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV  
APGEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFM  
WQQVFPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP  
SSKALICSFPSLQLEQATHQPIYLPRLGT

103/330  
MDILVPLLQLLVLLLTPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKREIFSQI  
KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV  
APGEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFM  
WQQVFPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP  
SSKALICSFPSLQLEQATHQPIYLPRLGT

**FIGURE 104**

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAGTTGCCTCATCGCAGG  
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA  
ACTAATATTTATATGACAGAAGAAAAAGATGTCATTCCGTAAAGTAAACATCATCATCTTGG  
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG  
TTAAGGAATGAGGTTACAGATTGAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA  
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATCTG  
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAAT  
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG  
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTGACCCTAAACTTTTGAAGGAA  
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTCTAC  
TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA  
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG  
AAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC  
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG  
CTCATTTAATCCTGGAGTTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATACTA  
ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT  
GGTAGCATCACAAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC  
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA  
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT  
ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTC AACCTAATCCGAAG  
ATATACCGAGATCTCAAACATAAAGTGAACAGAAATTTGAACTGTAAGCAAGCATTTCTCAG  
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA  
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAAACTGCCCTGTCTGGCAGTC  
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTCTTCTT  
ACTACAATGCTGAATGACTGGAAAGAAGAAGTATGCTAGTTGCTAGCTAGCTGGTACAGA  
TAATTCAAACTGCTGTTGGTTTTAATTTTGTAACTGTGGCCTGATCTGTAAATAAACTT  
ACATTTTTC

105/330

**FIGURE 105**

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR  
QEEIPVVIAASEDRLGGAIAAINSIOHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK  
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT  
ALKPGHAAAFSEDCDSASTKV VIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA  
NLTEWKRQINITNQLEKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTIDPMWNVRLGS  
SAGKRYSPQFVKA AKLLHWNGHLKPWGRTASYTDVWEKWIIPDPTGKFNLIRRYTEISNIK

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR  
QEEIPVVIAASEDRLGGAIAAINSIOHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK  
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT  
ALKPGHAAAFSEDCDSASTKV VIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA  
NLTEWKRQINITNQLEKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTIDPMWNVRLGS  
SAGKRYSPQFVKA AKLLHWNGHLKPWGRTASYTDVWEKWIIPDPTGKFNLIRRYTEISNIK



**FIGURE 107**

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG  
GAGCTAGAGGGCAAGTGTGCTCGGCCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC  
TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTTCGGGGCGCGGGCTGCA  
TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT  
TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT  
TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT  
CCGTAAAGTAAACATCATCATCTTGGTCCTGGGCTGTTGCTCTCTTCTTACTGGTTTTGCAC  
CATAACTTCCTCAGCTTGAGGCAGTTTGTAAAGGAATGAGGTTACAGATTCAGGAATTGTAG  
GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA  
GGAGATTCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA  
ACAGCATTCAGCACAACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA  
GCAGACCATCTCCGGTCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG  
TCAATTTTGACCCTAACTTTTGGGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC  
ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGG  
CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA  
CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTGAGCCTCTACTAAAGTTGT  
CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA  
TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC  
CTGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAAACTCAATGT  
AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTATCG  
TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT  
GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA  
TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA  
GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA  
CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG  
TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG  
GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT  
ATGTCTCCATCTGCCTTACCAAGTGTCTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA  
CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAACTGCTGTTGGTTTTAATTTT  
GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTTCAATAGGTAAAAAAGAAAAA  
AAAAA

**FIGURE 108**

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT  
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG  
TCTGTGCTGGTCTGAGGGTGCTGCCTGTCAATGGGGGCAGCCATCTCCCAGGGGGGCCCTCATC  
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCATCCTCTGCTG  
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC  
TGGCCCCTGTCCTGAGAAGGCCCCACCACCCCAGAAGCCCAGCCATGAAGGCAGCTACCTGC  
TGCAGCCCTGAAGGGCCCTGGCCTAGCCTGGAGCCCAGGACCTAAGTCCACCTCACCTAGAG  
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT  
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC  
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT  
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTCCCTAGGCTGAGCACC  
TCTAGGCCCTCTAGGTTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG  
GCTGGGGCCCCTCCCCTGGTCCTCCCAGTGTTTGCTGGATAATAAATGGAACCTATGGCTCTAA  
AAAAAAAAAAAAAAAAAAAA



109/330

**FIGURE 109**

MGA AISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH  
PRSPAMKAATCCSPEGPWPSLEPRT

109/330  
MGA AISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH  
PRSPAMKAATCCSPEGPWPSLEPRT

**FIGURE 110**

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA  
GTTCCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA  
CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGA  
TC**ATG**TCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC  
ATGATGTTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCATTGGTTAT  
TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC  
TCAGCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTA  
TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT  
GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC  
AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTG  
AGCCTGGGAAGTGCAGGAGCTGCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCCT  
TTCGGGCATTTCGGTACATGTGGTCTGACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA  
TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTACTTGTTATTTCAACAGAAGT  
AAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCA  
AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA  
TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA  
TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA  
TACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT  
TGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAATTTTTCTA  
GGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTACAATCG  
GGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGCCC  
ATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACCTTTTCTGTGTTTTGCTGTTGAT  
CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT  
CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA  
ATGAGGAGGGGAACAGAACTCCAGGCCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCT  
GGAAACATTTCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT  
AGTGAATTTTTTTTTTAAAGACCTAATAAACCCCTATTCTTCCTCAAAA

**FIGURE 111**

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTNDL  
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ  
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI  
LACQOMTIAGAVVTCYFNRSKNDDHPILSSLSILFFYHQGTVVKGSEFLISVVRIPRIIVM  
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKIL  
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH  
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN  
EEGTELQAIVR

111/330  
MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTNDL  
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ  
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI  
LACQOMTIAGAVVTCYFNRSKNDDHPILSSLSILFFYHQGTVVKGSEFLISVVRIPRIIVM  
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKIL  
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH  
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN  
EEGTELQAIVR



[illegible][illegible]

**FIGURE 114**

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGCGCTGAGAGGACACGAGCTCTA  
TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT  
 GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCCTACAGTTTTTCTCTGACTCTAAT  
 TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATCCAAAGAGTGGTTGAAG  
 TGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGAACAAACATT  
 CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA  
 GGCTGGATGGCCCTGTTCCGGGCTCTCCTGAGAATGGCTGAGGAGGCGGCCCCGAAAACCTCC  
 TCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGAACAGTGAACCTTACTTCATGGCGTG  
 AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC  
 CACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGTGGCCAGAGTGGCTTTGATGC  
 GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTTCGGCAACCACATTGATGTGCTCACTGGC  
 AAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT  
 GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCTTAGAGTATAACAAAG  
 CCATCCGGAACCTACACCCGCTTCGATGACTGGTACCTGTGGGTTTCAAGTGTACAAGGGGACT  
 GTGTCCATGCCAGTCTTCCAGTCCCTTGAGGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG  
 AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG  
 GGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA  
 CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT  
 CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGC GGAT  
 TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG  
 GCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACCAACTTCATCCACAACAATGG  
 GTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGCATCCTGGGGGCTGGGGGGTACA  
 TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG  
 GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTC  
 GAAATTTTCAGAAAAACACTGTTAGTTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC  
 TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCA  
 CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCTCT  
 AGACTCCTCATAACCACTGGATAATTTTTTTATTTTTATTTTTTTGAGGCTAAACTATAATA  
 AATTGCTTTTGGCTATCATAAAA

MPFRLLIPLGLLCALLPQHGGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE  
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLSSKKAGVEVEAGWPCSGPLLRMAEEAARKL  
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM  
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK  
AIRNYTRFDDWYLVWQMYKGTVSMPVFQSLEAYWPGQLQSLIGDIDNAMRTFLNYYTVWKQFG  
GLPEFYNI PQGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG  
FATIKDLRDHKL DNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY  
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPQT  
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

**FIGURE 116**

bioRxiv preprint doi: <https://doi.org/10.1101/2017.06.01.148088>; this version posted June 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.









120/330

**FIGURE 120**

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVNVASECGFTDQ  
HYRALQQLQRDLGPHHFENVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG  
AHPAFKYLAQTSGKEPTWNFWKYL VAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

120/330  
MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVNVASECGFTDQ  
HYRALQQLQRDLGPHHFENVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG  
AHPAFKYLAQTSGKEPTWNFWKYL VAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

**FIGURE 121**

CGGACGCGTGGGCGGGCCGGGACGCAGGGCAAAGCGAGGCC**ATG**GCTGTCTACGTCGGGATGC  
TGC GCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGCCCGGGCCGCCCTCTCT  
CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCAGAGAGGTTGGATCG  
CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTTCAGGGGTGCACCAAAAAGCATCTTA  
ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGGCC  
TTGGTCGTCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA  
AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC  
CTAACTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCCAGGCGGGCATCATTCTGGTG  
TCTGTGAACCCAGCCTACCAGGCTATGGAAGTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA  
GGCCCTTGTGTTCCCCAAGCAATTCAGACCCAGCAATACTACAACGTCTGAAGCAGATCT  
GTCCAGAAAGTGGAGAATGCCCAGCCAGGGGCTTGAAGAGTCAGAGGCTCCAGATCTGACC  
ACAGTCATCTCGGTGGATGCCCCTTTGCCGGGGACCTGCTCCTGGATGAAGTGGTGGCGGC  
TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCCTGTCCTGCCATG  
ACCCCATCAACATCCAGTTCACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCACCTCTCC  
CACTACAACATTGTCAACAACCTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC  
ACCAGAGCAGTTGCGGATGATCCTGCCCCAACCCCTGTACCATTGCCTGGGTTCGGTGGCAG  
GCACAATGATGTGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC  
AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT  
GTTCTGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG  
GTGTCAATTGCTGGGTCCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT  
ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTT  
CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAAATTATGCCTCACACGGAGG  
CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGC  
ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT  
GGATCAGGACAAGTGGTATTGGACAGGAGATGTCGCCACAATGAATGAGCAGGGCTTCTGCA  
AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG  
CTCGAGGACTTCTTTCACACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA  
CGATCGGATGGGGGAAGAGATTTGTGCCTGCATTTCGGCTGAAGGACGGGGAGGAGACCACGG  
TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC  
GTGTTTGTCAAAACTACCCCTCACCATTTTCAGGAAAGATCCAGAAATTCAAACTTCGAGA  
GCAGATGGAACGACATCTAAATCTG**TGA**ATAAAGCAGCAGGCCTGTCCTGGCCGGTTGGCTT  
GACTCTCTCCTGTGAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC  
TGAGCCAGGCACATCAAATGTCAAGGAATTGACTGAACGAACTAAGAGCTCCTGGATGGGTC  
CGGGAACCTCGCCTGGGCACAAGGTGCCAAAAGGCAGGCAGCCTGCCAGGCCCTCCCTCCTG  
TCCATCCCCCACATTCCCCTGTCTGTCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT  
GAAAAAAAAAAAAAAAAA

**FIGURE 122**

MAVYVGMLRLGRLCAGSSGVLGARAALSRSWQEARLQGVRFLSSREVDRMVSTPIGGLSYVQ  
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG  
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY  
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN  
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY  
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY  
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG  
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPOKTEEAVDQDKWYWTGDVAT  
MNEQGFCKIVGRSKDMIIRGGENIYPAELEDDFFHTHPKVQEVQVVGVKDDRMGEEICACIRL  
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

**Signal Peptide:**

amino acids 1-22

**Transmembrane Domains:**

amino acids 140-161, 213-229, 312-334

**Putative AMP-binding Domain Signature:**

amino acids 260-271

**N-myristoylation Sites:**

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277,  
314-319, 318-323, 379-384, 380-385, 409-413

**N-glycosylation Site:**

amino acids 282-285







125/330

**FIGURE 125**

MDPARKAGAQAAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT  
EAVGAVETIHGQFSLAVRGCGSLPGKNDRGLDLHGLLAFIQLQQCAQDRCAKLNLTSRAL  
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCDFGNVTLTAANVTV  
SLPVRGCVQDEFCTRDGVTGPGFTLSGCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTT  
VASTTSVTTSTSAFVRPTSTTKPMPAPTSQTPRQGEHEASRDEEPRLTGGAAGHQDRSMSG  
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

MDPARKAGAQAAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT  
EAVGAVETIHGQFSLAVRGCGSLPGKNDRGLDLHGLLAFIQLQQCAQDRCAKLNLTSRAL  
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCDFGNVTLTAANVTV  
SLPVRGCVQDEFCTRDGVTGPGFTLSGCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTT  
VASTTSVTTSTSAFVRPTSTTKPMPAPTSQTPRQGEHEASRDEEPRLTGGAAGHQDRSMSG  
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

**FIGURE 126**

[illegible]

127/330

**FIGURE 127**

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK  
CSFNQKPAPGDEEAQVENLITANATEPQKQORTEVQPSGGSLWNLRLLEPLDANVDA

127/330  
MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK  
CSFNQKPAPGDEEAQVENLITANATEPQKQORTEVQPSGGSLWNLRLLEPLDANVDA

**FIGURE 128**

AAACTTGACGCCATGAAAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT  
CCTCTGCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG  
CGTCACGACCCGAGGCCTTTAACACCCCGTTCTGAACATCGACAAATTGCGATCTGCGTTT  
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAAGGAACTTCCTTT  
CCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT  
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT  
CTTTCCTGCCTCAGGAACTCCAATAAAACATTTTCCATCCAAA

AACTTGACGCCATGAAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT  
CCTCTGCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG  
CGTCACGACCCGAGGCCTTTAACACCCCGTTCTGAACATCGACAAATTGCGATCTGCGTTT  
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAAGGAACTTCCTTT  
CCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT  
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT  
CTTTCCTGCCTCAGGAACTCCAATAAAACATTTTCCATCCAAA

129/330

**FIGURE 129**

MKIPVLPVAVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSFAKDE  
FLNWHALFESIKRKLPFLNWDAFPKLGKGLRSATPDAQ

129/330  
MKIPVLPVAVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSFAKDE  
FLNWHALFESIKRKLPFLNWDAFPKLGKGLRSATPDAQ

**FIGURE 130**

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCC**ATG**GGGGTGGAGATTGC  
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCAGGTTGTTT  
TTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC  
TTTGCCGGCCACTCA**TGA**GAGTGTTTTTGTGTAAAGTATTTTTTAGAATACTGTTGACTTCT  
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA  
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT  
AGGTTCAAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTATC  
TTCACTATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT  
TATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCACAAATGTGATTGTTAATTTAAA  
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT  
AAAAGTTTTGAGTATATTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

**FIGURE 131**

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

**FIGURE 132**

GGGGAATCTGCAGTAGGTCTGCCGGCGATGGAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG  
 GCTGCTGTTGTTTCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAATGGA  
 AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAC  
 TGCAGCTGCTACCATGGTGT CATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG  
 GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA  
 GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTGAGCACTTTATT  
 TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA  
 GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC  
 ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT  
 CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG  
 GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG  
 ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACCTTGTTGATGCAGAATACACCAAAAAC  
 CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT  
 GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA  
 AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC  
 TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAACAGATCTCTCCAATGTCCA  
 AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA  
 GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG  
 AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT  
 TCCCAAAATGTTGAAAACCTGAACTATAGTAGTCATCATAGGACCATAGTCCTCTTTGTGGCA  
 ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA  
 TCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGA  
 AAGATTTAAAATGTGTCTAATACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA  
 CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC  
 ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA  
 TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA  
 CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG  
 AAACCTTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG  
 TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTAAAACCATAAACTCTGTTACTCAG  
 GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT  
 TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA



**FIGURE 133**

MEWWASSPLRLWLLLFLLP  
SAQGRQKESGSKWKVFIDQ  
INRSLENYEPCCSSQNCSCY  
HGVIE  
EDLTPFRGGISRKMMAEV  
VRRKLGTHYQITKNRLYRE  
NDCMFPSRCSGVEHFILEV  
IGRLPD  
MEMVINVRDYPQVPKWME  
PAIPVFSFSKTSEYHDIMY  
PAWTFWEGGPAVWPIYPTG  
LGRWDL  
FREDLVRSAQWPWKKNST  
AYFRGSRTSPERDPLILLS  
SRKNPKLVDAEYTKNQAWK  
SMKDT  
LGKPAAKDVHLVDHCKYK  
YLFNFRGVAASFRFKHLF  
LCGSLVFHVGDEWLEFFYP  
QLKPWVH  
YIPVKTDLSNVQELLQFV  
KANDDVAQEIAERGSQFIR  
NHLQMDDITCYWENLLSEY  
SKFLSY  
NVTRRKGYDQIIPKMLK  
TEL

133/330  
MEWWASSPLRLWLLLFLLP  
SAQGRQKESGSKWKVFIDQ  
INRSLENYEPCCSSQNCSCY  
HGVIE  
EDLTPFRGGISRKMMAEV  
VRRKLGTHYQITKNRLYRE  
NDCMFPSRCSGVEHFILEV  
IGRLPD  
MEMVINVRDYPQVPKWME  
PAIPVFSFSKTSEYHDIMY  
PAWTFWEGGPAVWPIYPTG  
LGRWDL  
FREDLVRSAQWPWKKNST  
AYFRGSRTSPERDPLILLS  
SRKNPKLVDAEYTKNQAWK  
SMKDT  
LGKPAAKDVHLVDHCKYK  
YLFNFRGVAASFRFKHLF  
LCGSLVFHVGDEWLEFFYP  
QLKPWVH  
YIPVKTDLSNVQELLQFV  
KANDDVAQEIAERGSQFIR  
NHLQMDDITCYWENLLSEY  
SKFLSY  
NVTRRKGYDQIIPKMLK  
TEL



**FIGURE 135**

MAPALLLIPAAALASFILAFGTGVEFVRFSLRPLLGGIPESGGPDARQGWLAAALQDRSILAP  
LAWDLGLLLLFVGQHSLSMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV  
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP  
RALRLFSLRHPVCVELLTVLWVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR  
KLHLLSRPQDGEAE

**Signal sequence:**

amino acids 1-13

**Transmembrane domains:**

amino acids 58-76, 99-113, 141-159, 203-222

**N-myristoylation sites:**

amino acids 37-43, 42-48, 229-235





**FIGURE 138**

CCTCTGTCCACTGCTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT  
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA  
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA  
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC  
AGACTCTTTCAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT  
TCAATCCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC  
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA  
AAAAACATTGCAAACATGTGTCTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC  
AAGCCTGTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT  
CCTTCTGTGGAGACACGGTGGAGAACTTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT  
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT  
TTTCTCTACTAGTTATGTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAA

## **FIGURE 139**

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS  
IWDYGNNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQKGKPGGPPPKGLMYSVN  
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

### **Signal Peptide:**

amino acids 1-20

### **N-myristoylation Sites:**

amino acids 67-72, 118-123, 163-168

### **Flavodoxin protein homology:**

amino acids 156-174

139/330  
MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS  
IWDYGNNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQKGKPGGPPPKGLMYSVN  
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN  
Signal Peptide: amino acids 1-20  
N-myristoylation Sites: amino acids 67-72, 118-123, 163-168  
Flavodoxin protein homology: amino acids 156-174





**FIGURE 141**

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLFLFGTTEEEIQEICIETLRLY  
TRKKPNYELLEKEVEKRRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEK  
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN  
NRRSRSGTYSSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ  
SKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSRSGHGRHR

141/330  
MND  
SLR  
TNV  
FVR  
FQ  
PET  
IAC  
ACI  
YLA  
ARAL  
QI  
PL  
PTR  
PHW  
FLFL  
FGT  
TEEE  
IQE  
ICI  
ETLR  
LY  
TRK  
KPN  
YEL  
LEKE  
VEK  
RRK  
VAL  
QEAK  
LKAK  
GLNP  
DGT  
PAL  
STL  
GGF  
SPAS  
KPSS  
PREV  
KAEK  
SPIS  
INV  
KTV  
KKE  
PED  
RQQ  
ASK  
SPYN  
GVR  
KDS  
KRS  
RNS  
RSAS  
RSRS  
RTR  
SRSR  
SHTP  
RRH  
YN  
NRR  
SRSG  
TYSS  
RSR  
SHSE  
SPRR  
HHNH  
GSP  
HLKA  
KHTR  
DDL  
KSSN  
RHGH  
KRKK  
SRS  
RSQ  
SKSR  
DHSD  
AAKK  
HRHE  
RGH  
HRDR  
RRER  
SRSF  
ERSH  
KSKH  
HGGS  
RSRS  
SGH  
GRHR

**FIGURE 142**

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA  
TTTTTTGAANNTATTTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT  
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT  
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNTTATTGCTTACTGATTTTTTTTG  
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAGA  
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAACTGGTTTGTGTACATGCA  
AGCTTATAGTTGAAATATTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT  
GTTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA  
GATTCGGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTTGGTACTACAGAAGAGGAAATCC  
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAACTATGAATTACTG  
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT  
GAATCCGGATGGAACCTCAGCCCTTCAACCCTGGGTGGATTTTCTCC



**FIGURE 144**

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEELRALAGKPRPRGRKE  
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY  
YMLGPAKETNIAVFWCLLTVTFSIKMFLT VTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV  
VREETLELGLEPGLASMTQNL EPLLKKQGDWALPVAKLAIRVGLAVVGSVLGAFLT FPGLR  
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA  
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT  
VVS LQYLTPLILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAARI  
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLA GS

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEELRALAGKPRPRGRKE  
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY  
YMLGPAKETNIAVFWCLLTVTFSIKMFLT VTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV  
VREETLELGLEPGLASMTQNL EPLLKKQGDWALPVAKLAIRVGLAVVGSVLGAFLT FPGLR  
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA  
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT  
VVS LQYLTPLILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAARI  
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLA GS

**FIGURE 145**

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC  
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT  
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCC  
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC  
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC  
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT  
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTCTTGGT  
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTTCCTGACAGTGACACGGCTGTACTTCAGC  
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC  
CATGCTGGTGCAAGCG

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC  
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT  
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCC  
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC  
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC  
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT  
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTCTTGGT  
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTTCCTGACAGTGACACGGCTGTACTTCAGC  
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC  
CATGCTGGTGCAAGCG





[illegible][illegible]



149/330

**FIGURE 149**

MKILVAFLLVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT  
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE  
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

149/330  
MKILVAFLLVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT  
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE  
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

150/330

## **FIGURE 150**

GGCACGAGCCAGGAACTAGGAGGTTCCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC  
**ATG**GGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG  
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG  
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC  
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT  
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT  
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG  
CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA  
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG  
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG  
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC  
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA  
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG  
GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAGA**ATGAACC  
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCCTGTTCTGTTTGGGA  
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

150/330  
GGCACGAGCCAGGAACTAGGAGGTTCCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC  
ATG GGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG  
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG  
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC  
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT  
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT  
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG  
CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA  
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG  
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG  
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC  
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA  
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG  
GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATGTAGAATGAACC  
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCCTGTTCTGTTTGGGA  
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVPKGRWVLITCCAPQPPPPITYSL  
CGTKNIKVAKKVVKTHEPASFNLVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK  
PVSELNANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDGQVHLQQRPCRQPANFSFLP  
SQTSDWFWCQAANNANVQHSALTVPVPPGGDQKMEDWQGPLESPIALPLRSTRRLSEEEFG  
GFRIGNGEVRGRKAAAM

**Signal Peptide:**

amino acids 1-18

**N-glycosylation Sites:**

amino acids 86-89, 132-135, 181-184

**FIGURE 152**

GGTCCTTA**ATG**GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG  
CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT  
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA  
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA  
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT  
TACAGAGCAACTGCGTGACATTAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCCTGC  
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT  
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC  
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT  
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC  
CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCAC  
AGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA  
TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCAGC  
GTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT  
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCAGTGCCTTGATTCCTT  
TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC  
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC  
TCTTTTTGTTTTGGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT  
GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG  
AATTTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT  
TGTTCTGTACTGATATTTAAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

**FIGURE 153**

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFI  
HYDCGNKTVTPVSPGLGKKNVTTAWKAQNPNVLRVVDILTEQLRDIQLENYTPKEPLTLQAR  
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS  
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

### Important features:

**Signal peptide:**

amino acids 1-25

**Transmembrane domain:**

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

**FIGURE 154**

GGGAAAGCCATTTGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTG  
CCTTGGGCCTCACAATTTTCATTCTGTTTTCTGACTTTCAGTTATATAACCGTGGA**ATG**GAG  
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG  
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAAATCATGAACTCTGGCTGTTGATCAAAA  
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAGAAGCTAGCAGAAGAC  
TCAACCTGGCCTCCCATAAACAGGACAGATTATTCAAGGTGATGGCAAAAATGGATTCTACAT  
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAAACCTCAAATTGGGAGGCCAAC  
CCACAGAACAGCATTTCTGGGCCAGGCTG**TAA**TCAGAATTGTCGTCGTACATGCTCAACAGC  
ATTGCTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCCTACCTTT  
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC  
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTCCTGTAT  
CATCCTTTTCAATAAACTGTATTTCATTTTGAAAAAAAAAAAAAAAAAAAAAAAAA

155/330

**FIGURE 155**

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLA  
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL

Figure 155: Amino acid sequence of the protein. The sequence is shown in two lines, with the first line containing the first 20 amino acids and the second line containing the remaining 20 amino acids. The sequence is: MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLAEDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL.

**FIGURE 156**

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC**ATGAAG**  
 CTCTTGTGGCAGGTAAGTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTTCGTCTA  
 CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCC  
 AGAACTGCCCCCTCCGTTTGTCTGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCCGG  
 GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA  
 CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGCAGT  
 TGGGCAGGAAGTCCATCCGGCAGATTGAGGTGGGGCCCTTCAACGGCCTGGCCAGCCTCAAC  
 ACCCTGGAGCTGTTTCGACAAGTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC  
 CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCTCTTACGCCTTCA  
 ACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT  
 GAGGGAGCTTTTGGGGGCTGTTCAACCTCAAGTATCTGAACCTGGGCATGTGCAACATTAA  
 AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTGAGGGAACCACT  
 TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCATG  
 AACTCACAGGTCAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAGT  
 CAACTTGGCCCAATAACCTCTCTTCTTTGCCCATGACCTCTTTACCCCGCTGAGGTACC  
 TGGTGGAGTTGCATCTACACCACAACCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC  
 TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTGATGCTCCCAT  
 GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT  
 TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT  
 CGGACTCCCCCTATGTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC  
 CTCCCCGCCACCAAGGATCTCTGTCTCAACGACGGCACCTTGAACCTTTTCCACGTGCTGC  
 TTTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG  
 GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTACCACAGT  
 AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCTTA  
 CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTACAGACTACC  
 CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT  
 GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG  
 CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC  
 ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC  
 AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCACAATTC  
 ATGACCATATTAACATAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC  
 CTGGGGAACTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTACAGCCCA  
 TACCAAGGACAAGGTACAGGAACTCAAATA**TGACT**CCCCCTCCCCCAAAAACTTATAAAAT  
 GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA  
 TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAATAATTAAAGA  
 CAAAAAGTCAAAACA



**FIGURE 157**

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT  
RRGLSEVPQGIPSNTRYLNLMMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS  
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY  
ISEGAFEGLEFNLYLNLGMCNIKDMPNLTPLVGLEELEMGNHFPFIRPGSFHGLSSLKKLW  
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHNPWNCDLW  
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL  
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN  
ASAYLNVSTAELENSTSNYSFFTFTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ  
TTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS  
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE  
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

[illegible]



**FIGURE 160**

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA  
 AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAAC**ATGGGC**  
 TTCAACCTGACTTTCCACCTTTCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT  
 GACAGTGGTTGGGTGGGCCACCAGTAACCTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAG  
 CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACT  
 AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACCTGTCCTTCTGTGTCTCCTTACCTCAG  
 AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC  
 CCAAAGTGTCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC  
 CTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT  
 CCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT  
 TTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC  
 TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA  
 GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG  
 GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT  
 AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG  
 AATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG  
 ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACAGAGTCTGG  
 AGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA  
 TATCAACATCACAGTGGATTTCTGGTTTGGTGCA**TGA**CCCTGGATCTTTTGGTGATGTTTGG  
 AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA  
 AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCTTTTTGTATTTTCTTAGCAGAGCT  
 CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG  
 AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT  
 AAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATAT  
 TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT  
 CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCT  
 GTGAAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA  
 GGACACAGTGAACCTTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAG  
 CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCAGGTATGCCT  
 TCCAGTGATGCCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAATGA  
 TTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAATAATA  
 TGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAAA

## **FIGURE 161**

MGFNLTfHLSYKfRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKT  
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV  
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEgKKFNRAKLLNVGYLEALKEEN  
WDCFI FHDVDLVPENDFNLYKCEEHPKHLVVG RNSTGYRLRYSgyFGVTALSREQFFKVNG  
FSNNYWGwGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNNAERMKLLHQVSR  
VWRTDGLSSCSYKLVSVEHNPLYINITVDFWFGA

### **Important features:**

#### **Signal peptide:**

amino acids 1-27

#### **N-glycosylation sites:**

amino acids 4-7, 220-223 and 335-338

#### **Xylose isomerase proteins:**

amino acids 191-201



**FIGURE 163**

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA  
 LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF  
 VVLPITEFLVGDLVVTQNTSLPWPSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ  
 MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL  
 GPTLIQTFQKMTVTNLNFGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTRDPGD  
 YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV  
 ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

**Important features of the protein:****Signal peptide:**

amino acids 1-24

**Transmembrane domain:**

amino acids 339-362

**N-glycosylation sites.**

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

164/330

## **FIGURE 164**

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACTGAGCTCCCAGATCT  
GGGCCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCA  
CAACAGACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTG  
GATGCCCATGTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCT  
GCGGCTGCTGTATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAG**AACCTACCTGCCCTG  
CCCCGTCCCCTCCCTTCCCTTATTTATTCTGCTGCCCCAGAACATAGGTCTTGGAATAAAA  
TGGCTGGTTCTTTTGTTTTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AA

164/330



165/330

**FIGURE 165**

MALSSQIWAACLLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH  
FPICIFCCGCCHRSKCGMCCKT

165/330  
MALSSQIWAACLLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH  
FPICIFCCGCCHRSKCGMCCKT





**FIGURE 168**

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA  
CGTCTCCTCCAGGGATGGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC  
ACCTGGCAGGCCCAGGCTGTTCCACCATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGA  
CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG  
AAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCAGGAGACCTGGGAGGAC  
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT  
CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG  
GCTCCCGGGAGCTCTACATGAGGCACCTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG  
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCG  
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT  
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT  
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTG  
GAAGACTCTGCTCTTGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCCCTGAAAGTCCA  
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG  
CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG  
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT  
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTTGCAATGTGGAGACA  
TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA



170/330

## **FIGURE 170**

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAAT**ATG**GCTGGTTCCCCAACATGCCTCA  
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG  
GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC  
TATTGTCTGGACCTTCAACACAACCCCTCTTGTCAACATACAGCCAGAAGGGGGCACTATCA  
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG  
CTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT  
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG  
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG  
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC  
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT  
GCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT  
GAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCT  
CCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG  
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTGCGGAAACTCCTAACATATGCCCCCAT  
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA  
TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC  
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATC**TAG**ACAGCAGTG  
CACTCCCCTAAGTCTCTGCTCA



**FIGURE 172**

CTGGTTCCCCAACATGCCTCACCCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC  
TCTGGACCCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC  
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC  
AGCCAGAAGGGGGCCTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA  
GATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGT  
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG  
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG  
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT  
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG  
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC  
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCT  
GTGTCTCCTGTTGGTGCCCCCTCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC  
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGGAA  
ACTCCTAACATATGCCCCCATTTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA  
TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA  
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG  
AATGTTATCTAGACAGCAGTGCCTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAA





174/330

**FIGURE 174**

MKMLLLLCGLTLVCVHAEESSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ  
IHVLENSLVLVKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI  
NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

174/330  
MKMLLLLCGLTLVCVHAEESSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ  
IHVLENSLVLVKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI  
NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

**FIGURE 175**

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAA  
TGGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG  
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTC  
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAG  
AAAAAGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCA  
CAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTC  
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA  
CATTCATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG  
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCAC  
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCT  
TGTTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTC  
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTA  
TCAGTAGTTTGAAAAAAAAAA

176/330

**FIGURE 176**

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA  
IPATTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNA  
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL  
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

176/330  
MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA  
IPATTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNA  
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL  
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

177/330

**FIGURE 177**

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACC**ATG**AGGCT  
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT  
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC  
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG  
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAA**TAG**TGAA  
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC  
CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

177/330  
GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACC**ATG**AGGCT  
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT  
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC  
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG  
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAA**TAG**TGAA  
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC  
CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

178/330

**FIGURE 178**

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVN LQVAKLNPPPEALAAKLEV  
KHCTDQISFKKRLSLKKSWWK

178/330  
MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVN LQVAKLNPPPEALAAKLEV  
KHCTDQISFKKRLSLKKSWWK

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG  
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG  
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCCTCTCCCCCTGCAGCCCTGCCCCCTC  
GAACTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG  
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG  
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG  
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCAC  
TCATCACTCCAGGCTCTGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGA  
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC  
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAACTCG  
CCCCACCACCCCCTCA

180/330

**FIGURE 180**

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCK  
YKSSQKQHSPVPEKAIPLITPGSATTC

180/330  
M  
E  
R  
V  
T  
L  
A  
L  
L  
L  
L  
A  
G  
L  
T  
A  
L  
E  
A  
N  
D  
P  
F  
A  
N  
K  
D  
D  
P  
F  
Y  
Y  
D  
W  
K  
N  
L  
Q  
L  
S  
G  
L  
I  
C  
G  
G  
L  
L  
A  
I  
A  
G  
I  
A  
A  
V  
L  
S  
G  
K  
C  
K  
Y  
K  
S  
S  
Q  
K  
Q  
H  
S  
P  
V  
P  
E  
K  
A  
I  
P  
L  
I  
T  
P  
G  
S  
A  
T  
T  
C





**FIGURE 182**

MSLLSLPWLGLRPVAMSPWLLLLLVGWSLLARILAWTYAFYNNCRRLQCFPPKRNWFWG  
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF  
IRFLKPWLGEKILLSGGDKWSRHRRLTPAFHFENILKSYITIFNKSANIMLDKWQHMLASEGS  
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY  
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEDG  
KALSDDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW  
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW  
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT  
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

MSLLSLPWLGLRPVAMSPWLLLLLVGWSLLARILAWTYAFYNNCRRLQCFPPKRNWFWG  
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF  
IRFLKPWLGEKILLSGGDKWSRHRRLTPAFHFENILKSYITIFNKSANIMLDKWQHMLASEGS  
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY  
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEDG  
KALSDDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW  
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW  
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT  
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

**FIGURE 183**

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATCATGTATAAGCTGGCCTCCTGC  
TGTTTGCTTTTCACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA  
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAAGTCCGGAGGAGCTAGAAA  
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG  
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT  
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATACA  
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGTTGAAGTGAAATAAGCATCTGT  
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAAC  
AGTGTGGAGAAAAACTAGGCAAACCTACACCCTGTTTATTGTTACCTGGAAATAAATCCTCT  
ATGTTTTGCACAAAAAAAAAAAAAA



**FIGURE 185**

GAACATTTTGTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT  
GGGGTTGCTGGTTTAAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA  
CCACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCC**ATG**CCCCCTCCCC  
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT  
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTGAAGAAGCCACCA  
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA  
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT  
CAGGGGTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG  
GAAGAGGCCAAAGAGGCCCCAGCCGACAAG**TGA**TCGCCCAAGCCTTACTCACCTCTCTCT  
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA  
CAAGCTCAGGAGGCGAATAAATGTTCAAAGTGT

GAACATTTTGTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT  
GGGGTTGCTGGTTTAAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA  
CCACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCC**ATG**CCCCCTCCCC  
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT  
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTGAAGAAGCCACCA  
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA  
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT  
CAGGGGTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG  
GAAGAGGCCAAAGAGGCCCCAGCCGACAAG**TGA**TCGCCCAAGCCTTACTCACCTCTCTCT  
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA  
CAAGCTCAGGAGGCGAATAAATGTTCAAAGTGT

186/330

**FIGURE 186**

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG  
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG  
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO



188/330

**FIGURE 188**

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW  
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ  
AQQEAEELTPRPAGVVPGA

188/330  
MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW  
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ  
AQQEAEELTPRPAGVVPGA



**FIGURE 189**

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC**ATG**GCCAAG  
ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT  
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG  
TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG  
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACTGGGAGACTGGGGATGA  
CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAG  
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA  
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA  
GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCCTCCCCTTGGGGCTTTGTGGCAAAAATCCTA  
TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT  
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAAGTGGAGCGCCTTTGCTG  
CTGTTTCTCTGTCTGTGCTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTACAAAGTC  
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGGAATTATGGCTG  
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA  
ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAAC  
CCGAAGTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCAC  
CGTGGGTCTTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG  
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTG  
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA  
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG  
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC  
CTAAGGGATTCTTGGGTGCCACTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTACCCAC  
CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC  
TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

190/330

**FIGURE 190**

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP  
VSLDGDNTNTSTQEVVQYNWETGDDRF SFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR  
GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL  
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLTTTCLE  
LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

190/330  
MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP  
VSLDGDNTNTSTQEVVQYNWETGDDRF SFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR  
GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL  
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLTTTCLE  
LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

**FIGURE 191**

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG  
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA  
GCAGCTCCTGTGGACATTGCCATCCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG  
TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT  
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG  
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCCTCCTCCCAA  
GTCTGTTCTCTTATTGTCAACCTCAGCACAAACAGGCTGGCGCCAATGGCATTACAGAGAAAG  
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT  
AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTGAG  
GACATTCGCCCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT  
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTACT  
ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA  
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAA  
GGGAAGCAACAGGAACTTCTGCAACTGGTTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC  
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA  
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG  
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG  
AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC  
CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA



**FIGURE 193**

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTCGTGGAGCCAGGAGCGACGTCA  
 CCGCC**ATG**GCGAGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTGATGCTT  
 GGATGTGCCCTTCCAATATACAACAAATACTGGCCCCCTCTTTGTTCTATTTTTTTTACATCCTTTCACCTATTCC  
 ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACTTGCCATCTTTC  
 TTACAACGGGCATTGTCTGTCTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA  
 GCTTGTGCACTTGTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG  
 CAATGACGACTTCAGCTGGCAGCAGTGG**TGA**AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCATTT  
 GTTGGCCATTACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA  
 GGTGCTCCCTTCTCAGCTTTTATTGTAAGCATACTATTTTTCACAGAGACTTGCTGAAGGATTAAAGGATTTTCT  
 CTTTTGGAAGCTTGACTGATTTTCACTTATCTATAGTATGCTTTTTTGTGGTGTCTTGCTGAATTTAAATAT  
 TTATGTGTTTTTCTGTTAGGTTGATTTTTTTTGGAAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA  
 TTTGCATTGGTTAGGAATTCAGAATTCGCCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAATTTATT  
 TAGCCTCCATTATTACAAAAAATTATAAAAAATAAGTTTTTCAGTCAGTCAGGATGACATCACTCCCAATGTTATG  
 CAGACATACAGACGGTTGGCATACGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG  
 GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG  
 AAAATTATTTTATGGAATTGCTACAGAGGAGTGCTTTTTCTCTCTCAATTGTTAGAAGAATTTATGTTAACTTTA  
 AGGTAAGGGTGTA AAAACATTTTTGAGATAAGGTTTTATTTATGTTTATTATTGTTAGAGTGAGTTGCAATGT  
 GGGAAAGAAATGACATTGAAATTCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC  
 AACCTTTCATGTTTTACCTGTAAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC  
 ATCATATATGCCAGAAAACCTTCTCTGCTTCTCTCTTTTGACTTATTTGGTATGTTGTATATATTACATAAAA  
 TAACTTTTCAAATATAGTTTAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT  
 CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAACCTTGTTAGTCTTACAGATAAATTCATGCA  
 TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCCTTATTCTCTAAGGTTATATCATATGTAATTTAAAG  
 TATTTTTAAGACAAGTTTCTGTATACCTCTGAACTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT  
 CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT  
 ACCTGACCAAAAAATTTCCAGTAACCAGGCATGATCAATTTATAGTGGTTCGTTTACATCTAATAATTATCAGGA  
 CTTTTTTTCAGGAGTGGGTTATAAAAAACATTCAAGTTGGTCTGACAGTATTTTGTTAAGGATATTTGTTTGTATG  
 TTTATTCAGTATACCTTACATAAAAAATTATTTGCGCCATCAGCCAAAACCTCAGTAATCATGACAGCTGTCTGTTGT  
 TTTATGAAGTTTATTTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTTCTGCTTTTTTACTAAAGATGCCTAA  
 AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG  
 GCGTGTGGCTGGAGCCTTCCCCTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTTCAAGAGGAA  
 GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACTCTTTGTGCTTGTGATCTACTGGACTTT  
 TTTTTTGCAGGAAGTGCAATTCCTGCTCCTTCCCTATTTTCTGTTCTGGATGTGAGTGAGTGCACTGCTACTG  
 TTTTATCCACTTGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT  
 GTGTCTTTGACCTTGATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT  
 TTTCCATAGAATATGCACTGATACAACATTACCATTCTTCTATGAAAGAAAACCTTTTGATGATGAAACAATAA  
 AGATTTTAAATATCTATTTTAAAAA

# FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM  
 SNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFFLVFGSND  
 DFSWQQW

194/330  
 MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM  
 SNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFFLVFGSND  
 DFSWQQW

[illegible]















**FIGURE 202**

MTSKFILVSFILAALSSTTFSLQLDQQKVLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK  
 QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEETPIW  
 ITNQIRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY  
 WEDPDDMGHHLGPDSPMLGPFVISDIDKKLGYLIQMLKKAKLWNTLNLIITS DHGMTQCSEER  
 LIELDQYLDKDHHTLIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN  
 SRIQPIIAVADEGWHLQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD  
 LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYQTQSTILLPGSVKPAEYDQEGSYPYF  
 IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

**Signal Peptide:**

amino acids 1-22

**Transmembrane Domain:**

amino acids 429-452

**N-glycosylation sites:**

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

**Somatomedin B Domain:**

amino acids 69-85

**Sulfatase protein Region:**

amino acids 212-241

1000  
900  
800  
700  
600  
500  
400  
300  
200  
100  
0







**FIGURE 205**

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCCGCCAAC**AT**  
**G**CTCTGTCTGTGCCTGTACGTGCCGGTTCATCGGGGAAGCCCAGACCGAGTTCAGTACTTTG  
AGTCGAAGGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCCTCC  
CAGGAATTCTCCACCTACCGCCAGTGGGAAGCAGAAAAATTGTACAAGCTGGAGATAAGGACCT  
TGATGGGCAGCTAGACTTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA  
GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG  
CAGTCCCTGCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAG  
CATGGATAAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC  
ACCCCGTGGAAAACATCCCCGAGATCATCTCTACTGGAAGCATTCCACGATCTTTGATGTG  
GGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTG  
GAGACACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGG  
ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT  
GGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA  
CGTCTCAAATTTGCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCC  
TTGTTGGTAGTGACCGAGGAGACTCTGAGGATTCACGAGAGGCTTGTGGCAGGGTCTTGGCA  
GGGGCCATCGCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCG  
GAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGG  
TGGCCGCTTCTACAAAGGCTATGTCCCCAACATGCTGGGCATCATCCCCTATGCCGGCATC  
GACCTTGCAGTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTGAACAGCGC  
GGACCCCGGCGTGTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGG  
CCAGCTACCCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCT  
CCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAGGGGGCCTTCGGGCT  
GTACAGGGGGCTGGCCCCAACTTCATGAAGGTTCATCCAGCTGTGAGCATCAGCTACGTGG  
TCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCCGGT**G**ACGGGGGGAGGGCCGCCCG  
GCAGTGGACTCGCTGATCCTGGGCCGAGCCTGGGGTGTGCAGCCATCTCATTTCTGTGAATG  
TGCCAACACTAAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGAGGGAGGGT  
GGGGAGAGCTGGCAGGCCAGGGCTTGTCTGCTGACCCAGCAGACCCTCCTGTTGGTTCC  
AGCGAAGACCACAGGCATTCCTTAGGGTCCAGGGTCAGCAGGCTCCGGGCTCACATGTGTAA  
GGACAGGACATTTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCTTAGT  
TCTTCCATTTTACCCTTGCAGCCAGCTGTTGGCCACGGCCCCCTGCCCTCTGGTCTGCCGTGC  
ATCTCCCTGTGCCCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAG  
CCCACATCCCACCCCTCGTCCAATCCCATAATCCATGATGAAAGGTGAGGTACAGTGGCCT  
CCAGGCTGTACTTCCCAACCTACAGCATTGACGCCAACTTGGCTGTGAAGGAAGAGGAAAG  
GATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATGCT  
TGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCA  
TGGTGTCTGTAGCTGGCCTGGACCTGTGAGGATGGGGCCACCTCAGAACCAAACCTCACTG  
TCCCCACTGTGGCATGAGGGCAGTGGAGCACCATGTTTGAGGGCGAAGGGCAGAGCGTTTGT  
GTGTTCTGGGGAGGGGAAGGAAAAGGTGTTGGAGGCCTTAATTATGGACTGTTGGGAAAAGGG  
TTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG  
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTCCAACC  
CCAGCAGGGGGCGCAGCGGGACCAGCCCCACATTCACCTTGTGTCACTGCTTGGAACTATTT  
ATTTTGTATTTATTTGAACAGAGTTATGTCCTAACTATTTTTATAGATTTGTTAATTAATA  
GCTTGTCAATTTTCAAGTTCATTTTTTATTTCATATTTATGTTTCATGTTGATTGTACCTTCCC  
AAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGCCTTGGGCCGCTGCAGTCACAT  
CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAGGCAGCAGCCCTG  
GCTCCTTTCTTTGGCAGGTGGGGGAAGGGCTTGCCCCCAGCCTTAGGATTTTCAGGGTTTGA  
CTGGGGGCGTGGAGAGAGAGGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTT  
CTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT  
CACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCCTCC  
CAGCCTTCTGCTGCCCTTGCTTAACAATGCCGGCCAACCTGGCGACCTCACGGTTGCACTTCC  
ATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATT  
GTTATATGAACATATAACTGGAGTCGTCAAAAAGCAAATTAAGAAAGAATTGGACGTTAG  
AAGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAAAGAAAAA  
AA

**FIGURE 206**

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD  
 LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILK  
 SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW  
 WRHLVAGGGAGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI  
 NVLKIAPESAIAIKFMAYEQIKRLVGSQDETLEIRLHVAGSLAGAIQAQSSIYPMEVLKTRMAL  
 RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS  
 ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQAASIEGAPEVTMSSSLFKHILRTEGAFG  
 LYRGLAPNFMKVIPAUSISYVVYENLKITLGVQSR

**Important features:****Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 284-304, 339-360, 376-394

**Mitochondrial energy transfer proteins signature.**

amino acids 206-215, 300-309

**N-glycosylation site.**

amino acids 129-133, 169-173

**Elongation Factor-hand calcium-binding protein.**

amino acids 54-73, 85-104, 121-140

**FIGURE 207**

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCAT  
GGCTTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAG  
CAATTGCACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTC  
GCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAA  
ACTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCA  
AAGAAGGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTTTCAGAGGCCGGACAGCAGTGTTT  
GCTGATCAAGTGATAGTTGGCAATGCCCTCTTTCGGCTGAAAAACGTGCAACTCACAGATGC  
TGGCACCTACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATA  
AAACTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTG  
CGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCA  
GGGAGCCAACCTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCA  
TGAAGGTTGTGTCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAA  
AATGACATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAGGCGGAG  
TCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGTCTCTTCTTTCTTTGCCATCAGCT  
GGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAATGTGCCTTGGCCACAAAAAAG  
CATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTACCACCAGATATGACCTAG  
TTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCA  
AGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAATCTATCTTCAA  
GACATATTAGAAGTTGGGAAAATAATTCATGTGAAGTAGACAAGTGTGTTAAGAGTGATAAG  
TAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCT  
GGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGCTG  
TAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCA  
CAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGG  
GGCGGCTGCATTTTATAGTAATGGGTCAAATGATTCACCTTTTATGATGCTTCCAAAGGTGCCT  
TGGCTTCTCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAA  
ACAGAGCAGTCGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAA  
AA



**FIGURE 209**

1

210/330

**FIGURE 210**

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLLQEMKTLFLNTEYLMPFLLNQCSSLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGSRHTCVSSFPEPISP EWVRTRPFPI LPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLLQEMKTLFLNTEYLMPFLLNQCSSLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGSRHTCVSSFPEPISP EWVRTRPFPI LPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

**FIGURE 211**

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG  
 GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTGCGAAAAGATTCCGCAATAAACT  
 TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT  
 TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACC**ATGA**AGGGCATCCTCG  
 TTGCTGGTATCACTGCAGTGCTTGTTCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT  
 TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG  
 CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT  
 TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT  
 GCTGAAGAACACTTTTCATTTTGTAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG  
 CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG  
 AATCTAATGGAACCTTCCTGTCTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTGTC  
 TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC  
 CAACGTCAGTAACGCCACCTGTCAGTTCCCTGTCTGGTGAAAACAAGACTCTTGAGGAGTCA  
 TCTTTCGAAAGTTTGAGTGTGCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACT  
 TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGCCCTTGCCAGCCTCCTTCTTCG  
 GGGACTGCTGCCCC**TGA**GGTCCTGGGGCTGCACCTTGCCCAGCACCCCATTTCTGCTTCTCTG  
 AGGTCCAGAGCACCCCCTGCGGTGCTGACACCCTCTTTCCTGCTCTGCCCCGTTAACTGC  
 CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA  
 AAGCACTGGTTCATTCAGTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA







214/330

FIGURE 214

MVPAAGALLWVLLLNLGPRAGAQLTQTPTMQRVSLRFGGPMTRSYRSTARTGLPRKTRI  
ILEDENDAMADADRLAGPAAAELLAATVSTGFSRSSAINEDGSSEEGVVINAGKDSTSREL  
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP  
SPSPTAMPSPEDLRLVLPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC  
TYQQPCPNRLREECPLDTSLCTDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA  
FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100





**FIGURE 217**

[illegible]

**FIGURE 218**

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPESSTKETERKETKAEEL  
 DAEVLEVFPHTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN  
 TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV  
 RLINKFNSSSSSLEEKIAALFDLEYVHQMDNAQDLLSFGGLOVVINGLNSTEPLVKEYAAF  
 VLGAAFSSNPQVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL  
 GGLQVLRITLVQEKGTAVLAVRVVTLTYDLVTEKMFEEEEAEELTQEMSPEKLQYRQVHLLPG  
 LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS  
 LELQDGEDEGYFQELLGSVNSLLKELR

**Important features:****Signal peptide:**

amino acids 1-29

**Hypothetical YJL126w/YLR351c/yhcX family protein.**

amino acids 364-373

**N-glycosylation site.**

amino acids 193-197, 236-240

**N-myristoylation site.**

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

**Homologous region SLS1 protein.**

amino acids 68-340











**FIGURE 223**

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCCTTCCCTTTCCCCG  
GGGTCTGGGGTGACATTGCACGGGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC  
CCAGNTGGNGCGCCCTTCCCATTTCCTGTCCTGGTCAGGCCCCACCCCCCTTCCCACNTG  
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTTCGGCCCGGCCTTCG  
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA  
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC  
CGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC  
TTCTACAGGAGGTGTTCCGCTTTCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA  
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG  
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG  
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

**FIGURE 224**

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTCCCNTTCCCCGGGG  
TCTGGGGGTGACATTGCACCGCGCCCNCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCCC  
CAGNTGGCGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCTTCCCACCTGA  
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTTCGGGCCCGGCCTTC  
GCGCTTTTCTTGATCACTGTGGCTGGGGACCGCTTCGCGTTATCATCCTGGTCGCAGGGGC  
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA  
CCGACCGGTGAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC  
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGT  
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG  
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT  
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

**FIGURE 225**

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCCAGGAGGAGGCAG  
TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC  
TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG  
TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT  
AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT  
CCTTGTGGTTTCTCTACCTGGGGAAATAAGGTGCAGCGGCCATGGCTACAGCAAGACCCC  
CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT  
CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA  
GGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGGTGGATGACAGCAGCAGCCGCATCATCA  
ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC  
CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG  
GAAGAAAGTTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC  
AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCCTGGCTACTCCCACCCCTGGCCACTCT  
AACGACCTCATGCTCATCAAACCTGAACAGAAGAATTTCGTCCCCTAAAGATGTCAGACCCAT  
CAACGTCTCCTCTCATTTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA  
CCAAGAGCCCCCAAGTGCACCTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT  
CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA  
CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC  
TGCAGGGACTCGTGTCTTGGGGAGATTACCCTTGCGCCCGGCCAACAGACCGGGTGTCTAC  
ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAT  
CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTTCAG  
ACCCTCATTCCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCCTGACCCCATGTCT  
CCTGGACTCAGGGTCTGCTTCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG  
GAACAATTTCCAAAACCTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCAT  
CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA  
CTGAGAAGTGGAATAAAAAA









**FIGURE 229**

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT  
TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT  
CACCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG  
ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACTCGGCATCCAGAGCCC  
CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC  
TTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACC  
AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC  
CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC  
TGAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG  
AACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG  
GACTGTAAATATTTCTGCCTTAGTGAAACTCTACCATGCTGAAGATAAACAAACAAGAAGA  
CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT  
TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCCTTCTGAACTG  
TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG  
GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA  
TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGACT**TGA**ATTCGCC  
CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT  
TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCCTCGTG  
TTTCCTGTTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA  
AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG  
GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTTCAT  
GTCTTCCTTACACTTGGTGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC  
ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC  
AGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCAT  
CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG  
AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG  
TCTCCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG  
AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA  
TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG  
GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

230/330

**FIGURE 230**

MQAKYSSTRDMLDDDGDTTMSLSHQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL  
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE  
LYNKAGAHRCSPCTEQWKWHGDNICYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS  
QSYSEFFYSYWTGLLRPD SGKAWLWMDGTPFTSELFHIIIDVTS PRSRDCVAILNGMIFSKD  
CKELKRCVCERRAGMVKPESLHVP PETLGED

[illegible]



**FIGURE 232**

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGGGCTC  
 CGGGGATTTCGGCTCGGGCCGCTGGCTCTGCTCTGCGGGGAGGGAGCGGGCCCCGCGGGG  
 CCCGAGCCCTCCGGATCCGCCCCCTCCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT  
 CTGGGGGTTCGCCGGGGCCGGGGACCCGCGGTCCGGGGCGCC**ATG**CGGGCATCGCTGCTGCTG  
 TCGGTGCTGCGGGCCGCGAGGGCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCCTGAGCCT  
 GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGGCCAGGCCCCGCCCCAACCTGGAGACTCTG  
 AGCTGCCGCCGCGCGGCAACACCAACGCGGCGCGCCGGCCCAACTCGGTGCAGCCCCGGAGCG  
 GAGCGCGAGAAGCCCCGGGGCCGGCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC  
 CTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTCAGGACCCGCTACATCAGCA  
 CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC  
 ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC  
 GGGCGCACGGGGCCGCGGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC  
 CCATTGGACACCTGCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGAC  
 TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG  
 CCACCTCAGCCTGGCCTCCGCGCCCACTGTACCTGGGCGGGCCCCAGGACTTCATCGGCG  
 GAGAGCCCACCCCCGGCCGCTACTGCCACGGAGGCTTTGGGGTGCTGCTGTGCGCATGCTG  
 CTGCAACAACCTGCGCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCGCCCTGA  
 CGAGTGGCTGGGTGCTGCACTTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG  
 GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGCAGGAGGGGGACCCTCAT  
 TTCCGAAGTGCCCTGACAGCCACCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA  
 AGCTTTGCCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA  
 TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCCGTGGGTATT  
 CCAGCACCATCCCGCCCCGCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA  
 GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCACTGCGTGGGGCTGACCGGGCTG  
 ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCTTG  
 CGGCTCCAGAAGCAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCCGGGTATGGA  
 ATACACGCTGGACTTGCACTGGAGGCACTGACCCCCAGGGAGGCCGCGGGCCCCCTCACTC  
 GCCGAGTGCACTGCTCCGGCCGCTGAGCCGCTGGAGATCTTGCTGTGCCCTATGTCACT  
 GAGGCCTCACGTCTCACTGTGCTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCCTGG  
 CTTCTTGAGAGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC  
 TGCTGCTACTGTATGAGCCGCGCCAGGCCAGCGCGTGGCCCATGCAGATGTCTTCGACCT  
 GTCAAGGCCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGTGCCATGGCTCAG  
 TGTGCAGACAGCCGCACCCTCACCCTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC  
 TGGACACACTGTTCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC  
 TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA  
 CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCAGAGCTGGGCGGTGACACTGGCCGCT  
 TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACTCCGACTACGTGGCAGCCCGTGGG  
 CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGGATGTGTACGAGCT  
 GTTCTTCCACTTCTCCAGTCTGCATGTGCTGCGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT  
 ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC  
 GTGCTTGAGGGCCTCGGCTCCCGAACCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG  
 CAACAGCACCT**TGA**CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCACCCCACTT  
 CTCCCCCAAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC  
 AAGCTGGCCCACTGGTCCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG  
 GGACGTGCCCCCAGAGCCACCCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT  
 GCTGATTTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC  
 CTCTGGGCCCTGGGGGCTGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGGAGCTGAGGAGGGG  
 GCATCTCCCAACTTCTCCCTTTTGACCCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA  
 AGTGTGGAAAAA

**FIGURE 233**

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELPPRGNTNAARRP  
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAACKAVRTRYISTELGIRQRLLVAVL  
TSQTTLP TLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERP IGH LHLALRH LLE  
QHGD DFDWFFLV P D TTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG  
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGR CIL DATGVGCTGDHEGVHYSHLELSPGEP  
VQEGDPHFRSALTAHPVRDPVHMYQLHKAFARAELE RTYQEIQELQWEIQNTSHLAVDGDRA  
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN  
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI  
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA  
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL  
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGP GPPELGRDTGRFDRQAASEACFYNS  
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL  
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST



235/330

**FIGURE 235**

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE  
DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER  
LVKYREANGLPIMESNCFDPSKIQLPED

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE  
DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER  
LVKYREANGLPIMESNCFDPSKIQLPED







**FIGURE 238**

MPLALLVLLLLGPGGWCLAEPFRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY  
RLFPKALGQLISKYSLRELHLSFTQGFWRTYWGPPFLQAPSGAELWVWFQDTVTDVDSWK  
ELSNVLSGIFCASLNFIDSTNTVTPTASF KPLGLANDTDHYFLRYAVLPREVVCTENLTPWK  
KLLPCSSKAGLSVLLKADRLFHTSYHSQA VHIRPVCRNARCTSSISWELRQTL SVVFDAFITG  
QGKKDWSLFRMF SRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT  
YAIYDLLDTAMINNSRNLNIQLKWKRP PENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH  
PYRAFPVLLLLDTPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS  
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSD  
GSNYFVRLYTEPLL VNLPTPDFSMPYNVICLTCTVVAVCYGSFYNNLLTRTFHIEEPRTGGLA  
KRLANLIRRARGVPPL

**FIGURE 239**

CAACATGGGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG  
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC  
TGCTTCAAGTCCGATCCTCCCCAGTGTCACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG  
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTGTGAAGGAACTGGAAGAAGGAGGAA  
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC  
TCCTCCTCTACCAGGTGTCCTCAGAATGATGCTGGGTCCTTTCTACCTCTGGGGGTCACTC  
TCACTTGGCACCTGCCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC  
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAAG  
TCCAGACCCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT  
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC  
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAACTGGTCCTCACCAGATG  
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT  
TTGCTGTTTATAATCCAA

240/330

## **FIGURE 240**

MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC  
YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

### **Signal sequence:**

amino acids 1-19

### **N-myristoylation sites:**

amino acids 23-29, 27-33, 32-38, 102-108

### **WAP-type 'four-disulfide core' domain signature:**

amino acids 49-63

MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC  
YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK



**FIGURE 242**

MRSCLWRCRHLISQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP  
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW  
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNNGQTRKLTASRTVSEKHQG  
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN  
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF  
NQSEWDRLEHFAPPFGEFMEFNYSLVQKVVTREPPVPQQQLLASLPAGSLRCITCAVVGNGG  
ILNNSHMGQEIDSHDYVFRSLGALIKGYEQDVGTRTSFYGFATAFSLTQSLLILGNRGFKNVP  
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL  
RYMKNRFLRSKTLTGAWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSW  
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

**Cytoplasmic Domain:**

amino acids 1-10

**Type II Transmembrane Domain:**

amino acids 11-35

**Lumenal catalytic Domain:**

amino acids 36-600

**Ribonucleotide Reductase small subunit Signature:**

amino acids 481-496

**N-glycosylation Sites:**

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

**FIGURE 243**

CG**ATG**CGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG  
GAGCAGCGAGTGGAAATTGTTCCCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAAACACCT  
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA  
CCAGAGACTGCACCATTCTCCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT  
GTCAGTCCAGTGTGCATGGAGGATAAG**TGA**GCAGACCGTACAGGAGCAGCACACCAGGAGCC  
ATGAGAAGTGCCTTGGAAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG  
ACAAGAGATTTATTTTTGCAGACAGACTCTTCCATAAGTCCTTTGAGTTTTGTATGTTGTTG  
ACAGTTTGCAGATATATATTCGATAAATCAGTGTACTTGACAGTGTTATCTGTCACTTATTT

CGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG  
GAGCAGCGAGTGGAAATTGTTCCCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAAACACCT  
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA  
CCAGAGACTGCACCATTCTCCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT  
GTCAGTCCAGTGTGCATGGAGGATAAGTGAAGCAGACCGTACAGGAGCAGCACACCAGGAGCC  
ATGAGAAGTGCCTTGGAAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG  
ACAAGAGATTTATTTTTGCAGACAGACTCTTCCATAAGTCCTTTGAGTTTTGTATGTTGTTG  
ACAGTTTGCAGATATATATTCGATAAATCAGTGTACTTGACAGTGTTATCTGTCACTTATTT

244/330

## **FIGURE 244**

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFRLYHNT  
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFRLYHNT  
RDCTIPAYYKRCARLLTRLAVSPVCMEDK



**FIGURE 245**

GGGCTGGGCCCCGCCGAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG  
CCCGACCCCGGCCGCGCCCAGCCCCACCATGCCACCCGCGGGGCTCCGCCGGGCCGCGCCG  
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT  
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT  
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG  
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT  
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTTCTGCTACCTGT  
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTGAAGGCCAGGAGATTCCAATGACAGGCATC  
CCAGTGCAGCCAGTATACCCATACCCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC  
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC  
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC  
TGAGGAACCAGCCATGTCTCTGCTGCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT  
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA  
GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA  
ACTATGAGGGGTTGGGGGGAGGGCTTGGAATTATGGGCTATTTTTTACTGGGGGCAAGGGAGG  
GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG  
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT  
TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG  
CTAGATTAAAGCTGTAAAGACAAA





**FIGURE 248**

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTFPFWNKIRLPEYVIPVH  
YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE  
HPPQEQIALLAPEPLLVLGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA  
ARMAFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA  
FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA  
IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL  
WLNEGFAKFMEFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD  
DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGM  
FCSRSQHSSSSSHWHQEGVDVKTMNTWTLQRGFPLITITVGRNVHMKQEHYMKGSDGAPD  
TGYLWHVPLTFITSKSNMVHRELLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL  
TGLLKGTHTAVSSNDRASLINNAFQLVSIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP  
MYKLMEKRD MNVETQFKAFLIRLLRDLIDKQTTWDEGSVSEQMLRSELLLLACVHNYQPCV  
QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALC  
RTQNKEKLQWLLDESFKGDKIKTQEFQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS  
SSIAHMVMGTTNQFSTRTRLEE VKGFFSSLKENGSQLRCVQQT IETIEENIGWMDKNFDKIR  
VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360



250/330

**FIGURE 250**

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI  
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP  
PADPGSLRCFVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCN  
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL  
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN  
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC  
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESITWGVGLALAPALWWGVVC  
PSC

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI  
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP  
PADPGSLRCFVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCN  
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL  
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN  
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC  
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESITWGVGLALAPALWWGVVC  
PSC

**FIGURE 251**

GCGACGGGCAGGACGCCCCGTTTCGCCTAGCGCGTGCTCAGGAGTTGGTGTCTGCCTGCGCT  
CAGG**ATG**AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG  
CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG  
CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCCGGACGGCCTGGAAGAGTCG  
GCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTGCTCAT  
GGAAAAATTGGTCCCATTTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC  
TGGTCCTAATGGAGAACCAGGCCTCCCATGTGAGTGCAGCCAGCTGCGCAAGGCCATCGGGG  
AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTGCGC  
GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA  
CGCCCAGCTGTCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA  
ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCCGTGTCTTCATCGGCATCAAC  
GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA  
GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT  
CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG  
GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTGGGGGGCCCCACATGTCCCTGCAGGGTT  
GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG  
GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG  
AAAGTGTTCTTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA  
ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC  
TATACAATAAAATCTTTAAGTAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

**FIGURE 252**

M R G N L A L V G V L I S L A F L S L L P S G H P Q P A G D D A C S V Q I L V P G L K G D A G E K G D K G A P G R P G R V G  
P T G E K G D M G D K G Q K G S V G R H G K I G P I G S K G E K G D S G D I G P P G P N G E P G L P C E C S Q L R K A I G E  
M D N Q V S Q L T S E L K F I K N A V A G V R E T E S K I Y L L V K E E K R Y A D A Q L S C Q G R G G T L S M P K D E A A N  
G L M A A Y L A Q A G L A R V F I G I N D L E K E G A F V Y S D H S P M R T F N K W R S G E P N N A Y D E E D C V E M V A S  
G G W N D V A C H T T M Y F M C E F D K E N M

MRGNLALVGV  
LISLAFLLS  
LLPSGHPQ  
PAGDDACSV  
QILVPGLK  
GDAGEKGD  
KGAPGRPGR  
VPTGEKGD  
MGMGDKG  
QKGSVGRH  
GKIGPIGSK  
GEGKGDSD  
IGPPGPNG  
EPGLPCECS  
QLRKAIGE  
MDNQVSQ  
LTSELKF  
IKNAVAGV  
RETESKI  
YLLVKEEK  
RYAD  
AQLSCQGR  
GGTLSMP  
KDEAAN  
GLMAAYL  
AQAGLAR  
VFIGIND  
LEKEGA  
FVYS  
DHSPMRT  
FNK  
WRS  
GEPNNAY  
DEEDC  
VEMVAS  
GGW  
NDVACH  
TTMY  
FMCE  
FDKENM



**FIGURE 253**

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG  
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC  
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACCA**ATG**AGGATCATGCTGCTATTACAGCCAT  
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG  
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC  
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA  
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA  
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCCTCGGCCC  
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC  
TTTAT**TAA**GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT  
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC  
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT  
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA  
TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT  
CCTACATTAAAAATATAATGTCTCTCTCTATTCTCAACAATAAAGGATTTTGCATATGAA  
AA

**FIGURE 254**

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK  
ALSQASTDPKESTSPEKRDMDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK  
SSLGTEEQRPL

**Important features:****Signal peptide:**

amino acids 1-18

**Tyrosine kinase phosphorylation site.**

amino acids 36-45

**N-myristoylation site.**

amino acids 33-39, 59-65

**Amidation site.**

amino acids 90-94

**Leucine zipper pattern.**

amino acids 43-65

**Tachykinin family signature.**

amino acids 86-92

**FIGURE 255**

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCCTGTGCCTGCTGTGCC  
CGCGCTGTGCGCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG  
GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC  
CCATAGTTGCTGCAGGAGTGGAGCC**ATG**AGCTGCGTCCTGGGTGGTGTATCCCCCTTGGGGC  
TGCTGTTCCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTCTTAGAGGAG  
CTGCTCAGCAAATACCAGCACAAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA  
GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT  
CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC  
CTGGGCCACCGAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCCTGTGATGG  
CTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTTGTTTGTGTT  
TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG  
AAACCT**TAG**ACTCCCGGGGTAAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACCTACAG  
GCATGCACCATGGTGCCCGAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT  
TGCCCAGGCTGGTCTTGAACCTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG  
CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA  
ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC  
TCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGCCAGGG  
CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC  
CTTAGCCCCGTGTGAGCCTCACTTTCCACTTGGAGAGTCCTTCCTCGCGTGGTTGCCATGACT  
GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCTAGG  
CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT  
GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG  
TGGAACCTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG  
GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT  
TTTAAACAAACCCACGTCCCAGCCTGGGTAAACATGGTAAAGCCCCGTCTCTACAAAAAATC  
CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG  
GTGGAGGTGGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC  
AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

## **FIGURE 256**

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILML  
HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR  
GGSRLCSVLEFVCFETGSHSATDAGVQWHNRHALKP

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **N-glycosylation site.**

amino acids 27-31, 41-45

#### **N-myristoylation site.**

amino acids 126-132, 140-146

#### **Amidation site.**

amino acids 85-89



**FIGURE 258**

MSGGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTS  
VTLHHARSQHHVVCNT

MSGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTS  
VTLHHARSQHHVVCNT

**FIGURE 259**

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT  
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAAACAATTAAGTGTAGGATTGCAGT  
T**ATG**ATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCTCCACTGTGTTCTGTGTGC  
TATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTTGAAAATGAATGCTTAGTGGATCTG  
TGCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTTCCCCTCCC  
CTCCGATTGTTCTAAAT**TAA**ATTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTG  
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCCCTGAAAGCCAATGGAAATACTTTTTTT  
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA  
GTAAAATAGAAACCTGTGTTTATTCTCAGGTATTTTAGAAACAACAGCCATCATTTTATTTT  
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT  
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCATTTTCTC  
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA  
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAAATTGATATTAATAAAACCAGA  
ATAAAAGTTCATATCTACCC







**FIGURE 262**

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE  
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL  
QEAEKKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNPKVYLLIGSRNNTV  
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED  
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKIEPGTLGVEHSWDT  
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH  
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE  
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL  
QEAEKKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNPKVYLLIGSRNNTV  
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED  
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKIEPGTLGVEHSWDT  
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH  
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

[illegible]

**FIGURE 264**

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK  
 QYQRIRKEKPQQHNFTHRLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ  
 REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPPK  
 HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTAAVSPCGRFVASCG  
 FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV  
 EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH  
 GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ  
 AQETLKSLGALKK

**Important features:****Signal peptide:**

amino acids 1-25

**N-glycosylation site.**

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

**Beta-transducin family Trp-Asp repeat protein.**

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

**FIGURE 265**

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG  
CAGTGTTTTGCCTTCACCCCAAGTGACCATGAGAGGTGCCACGCGAGTCTCAATCATGCTCC  
TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT  
GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT  
GGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA  
AGCACCACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCTGGACGGCAGGTAC  
CGCTGCTCCATGGACTTGAAGAACATCAATTTTTAGGCGCTTGCCTGGTCTCAGGATACCCA  
CCATCCTTTTCCTGAGCACAGCCTGGATTTTTTATTTCTGCCATGAAACCCAGCTCCCATGAC  
TCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAG  
GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG  
CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTA  
AATGGCAGAAAGGACATTCCCCCTCCCCCTCCCCAGGTGACCTGCTCTCTTTCTGGGCCCTG  
CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTGGGT  
GCATTGCTCAGAGTCCCAGGTCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG  
ACCAATTTGTGGGTAGTTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC  
TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA  
GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA  
CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT  
TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA  
TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACCTCTTTCCA  
AGGATCAGCCCTGAGAGCAGGTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG  
TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTACAGACCAGGGAGG  
CAACTACACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAA

## **FIGURE 266**

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHP  
GSHKVPFFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

### **Signal peptide:**

amino acids 1-19

### **Tyrosine kinase phosphorylation site:**

amino acids 88-95

### **N-myristoylation sites:**

amino acids 33-39, 35-41, 46-52



**FIGURE 268**

MSFLQDPSFFTGMWSIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKEPR  
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF  
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEFGFILGGVFV  
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

MSFLQDPSFFTGMWSIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKEPR  
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF  
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEFGFILGGVFV  
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK



**FIGURE 269**

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCG  
GGCCAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA  
AGCCCCCTTCCTCGGGCGCTGCCAACCCGCCACCCAGCCC**ATG**GCGAACCCCGGGCTGGGGCTG  
CTTCTGGCGCTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACA  
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG  
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGCTCTTCTCCCTCTTGGCTGCCTTG  
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG  
CACCTACCGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG  
ACTCCAAGGAGACGGTGCAGGGCTGCCCTGCCCATC**TAG**GTCCCCCTCTCCTGCATCTGTCTCC  
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG  
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGG  
GCTATTCACTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

270/330

**FIGURE 270**

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSSDGNLRPEAITAIIV  
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSSDGNLRPEAITAIIV  
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

**FIGURE 271**

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT  
TTGGGATTTTAATTTTCAAACACAGCAGAAATGACATTTTTTCTGTCACTATTATTATTGTTG  
GTATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCT  
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC  
CTACCAAAGCTGTCAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA  
GGGTTAATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATG  
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATA  
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT  
TTATTAATTTTTTAAAATCAGTAACTGATTTTATCACTGGCTATGTGCTTAGATCTACAGGAGA  
TCATATAATTTGATACAAATAAAAGAAAAGTGTCTCTCCCCTTACAGAATTGACATTTTAA  
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA  
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

272/330

## **FIGURE 272**

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK  
GIVKGRNLDSRGLILGAEAWGRGVKKNT

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK  
GIVKGRNLDSRGLILGAEAWGRGVKKNT



**FIGURE 274**

MGLFRGFVFLLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTY  
 LFEATEKRFFFKNVSLIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC  
 GEKGEYIHFTPDL LLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR  
 CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVE  
 FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLV  
 LDKSGSMGGKDRLNRMNQAAKHFL LQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM  
 AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVH  
 FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT  
 LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG  
 TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG  
 YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNDBGVYSRYFTAYTENGRYSLKVRAGH  
 GANTARLKL RPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRTASGGAFVVSQV  
 PSLPLPDQYPPSQITDL DATVHEDKII LTWTAPGDNFVGVQRYIIRISASILDLRDSFDD  
 ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIP  
 QANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

**Signal peptide:**

amino acids 1-21

**Putative transmembrane domains:**

amino acids 284-300, 617-633

**Leucine zipper pattern.**

amino acids 469-491, 476-498

**N-glycosylation site.**

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,  
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

**FIGURE 275**

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG  
 GGCAGGGGTGACAAACAGGTGTCTCTTTTGTATCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAG  
 ACGCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACC  
 CCCAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTT  
 GTGGTTGGAGGAGAGAACCCTTTGTGGGGCTGCGTTCTCTTAGCAGTGCTCAGAAGTGACTTGCCTGA  
 GGGTGGACCGAGAAGAAAGGAAAGGTCCTTCTTGGCTGTGGCTGCACATCAGGAAGGCTGTGATGGG  
 AATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAAA  
 GTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATG  
 CAAGCAGCTCCGGGGGCCCAACGCATGCTTCTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGG  
 GCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTCCGCC  
 GGGGGCTGCTTGCCTGGATTTCCTGGGTGGTGGTTTGTCTGGTGCTCCTCTGCTGTGCTATCTCTGT  
 CCTGTACATGTTGGCCTGCACCCCAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGCCAACAGC  
 CCCACGGGGAAGGAGGGGTACCAGGCCGCTTTCAGGAGTGGGAGGAGCAGCACCGCAACTACGTGA  
 GCAGCCTGAAGCGGCGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAA  
 TGGGCAGTACCAAGCCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAG  
 GCCGACCTCCTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG  
 CCACAGAGTATGCAGCAGTGCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGG  
 CTTTACCCCGCACCCCGAGGAGAACCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATT  
 GAATCAGCCTTGGAGACCCTGAACAATCCTGCAGAGAACAGCCCCAATCACCCTCTTACACGGCCT  
 CTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTTCAA  
 AGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTACAGCCCCATCATGAAAGTG  
 AAAAATGAAAAGCTCAACATGGCCCAACACGCTTATCAATGTTATCGTGCTCTAGCAAAAAGGGTGG  
 ACAAGTTCCGGCAGTTTCATGCAGAATTTTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCT  
 CACTGTTGTTTACTTTGGGAAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAA  
 GCTGCCAATTCAGGAACCTTACCTTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTG  
 ATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGCTCCTTCTCTTTTCTGTGATGTGGACATCTACTT  
 CACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTTTATCCAGTT  
 CTTTTAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAGC  
 AGCTGGTCATAAAGAAGGAAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTGAGTATCG  
 GTCAGACTTCATCAATATAGGTGGGTTTGTATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCAC  
 CTTTATCGCAAGTATCTCCACAGCAACCTCATAGTGGTACGGACGCTGTGCGAGGACTCTTCCACC  
 TCTGGCATGAGAAGCGCTGCATGGACGAGCTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAA  
 GGCCATGAACGAGGCATCCACGGCCAGCTGGGCTGCTGGTTCAGGCACGAGATAGAGGCTCAC  
 CTTGCAAAACAGAAACAGAAAGACAAGTAGCAAAAAACATGAAGTCCCAGAGAAGGATTTGGGAGA  
 CACTTTTTCTTTCTTTTGTCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAGGACG  
 ACAAAGAATTTGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTCTGTTGGGCTTTTTTAC  
 AACAGAAATCAAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACA  
 AAGGCAGAAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTTGATGGTGTTTACAATACACT  
 GAGACCTGTTGTTTTGTGTGCTCATTGAAATATTCATGATTTAAGAGCAGTTTTGTAAAAAATTCAT  
 TAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTTCTAGG  
 AATGCTAAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACTAGTGAGTACATTAAGTA  
 AAATAAATGGAACAGAAAAGAAAAGAAACCATAAATATCGTGTCATATTTTCCCCAAGATTAACCA  
 AAAATAATCTGCTTATCTTTTTGGTGTCTTTTAACTGTCTCCGTTTTTTTTCTTTTATTTAAAAAT  
 GCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGA  
 GCACAAGTTGGCCTACATTTTTATATTTTTTAAGAAGATACCTTTGAGATGCATTATGAGAACTTTCA  
 GTTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAAT  
 GTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAA  
 GAGTATTTTCTGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAA  
 AAGGAAACTCATTACAGCTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACATTTTCTCCTCA  
 GAAGTAGGGACCGCTTTCTTACCTGTTTAAATAAACCAGTATACCGTGTGAACCAACATCTCT  
 TTTCAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATAT  
 ATATATATATATTGTGAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGCTACAT  
 GTTATCCACCCAGGCCAGGTGGAAGTAACTGAATTTATTTTTAAATTAAGCAGTTCTACTCAATCA  
 CCAAGATGCTTCTGAAAATTGCATTTTATTACCATTTCAAACTATTTTTTAAAAATAAATAACAGTTA  
 ACATAGAGTGGTTTCTTCAATTCATGTGAAAATTATTAGCCAGCACCAGATGCATGAGCTAATTATCT  
 CTTTGAGTCTTGTCTTGTGTTGCTCACGTAATACTCATTTGTTTAAAGCTTCAAGAATTTCAAGC  
 TGTGGTGTGTTAAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTTAAACAC  
 AGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTTGTGGATATGAA

**FIGURE 276**

MMMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ  
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL  
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIES  
ALETNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP  
MKVKNEKLNMAN TLIN VIVPLAKRVDKFRQFMQNFREMCIEQDGRVH LTVVYFGKEEINEVK  
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLN TCR  
LNTQPGKKVFYPVLFSQYNPGI IYGHHD AVPPLEQQ LVIKKETGFWRDFGFGMTCQYRSDFI  
NIGGFDL DIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTP EQYKMCMQS  
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

MMMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ  
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL  
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIES  
ALETNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP  
MKVKNEKLNMAN TLIN VIVPLAKRVDKFRQFMQNFREMCIEQDGRVH LTVVYFGKEEINEVK  
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLN TCR  
LNTQPGKKVFYPVLFSQYNPGI IYGHHD AVPPLEQQ LVIKKETGFWRDFGFGMTCQYRSDFI  
NIGGFDL DIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTP EQYKMCMQS  
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT







**FIGURE 279**

AACTCAAACCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG  
GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATGT**  
ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG  
CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC  
TCGGTTAAATGCACTTTCTCCAGCTTTGCCCTGTGGGTGATGCTCTAACAGTGACCTGGA  
ATTTTCGTCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC  
CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA  
TGCTCCATCCTTCTCTGGAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA  
AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA  
CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT  
AATAGTAATTGTAGTGGTCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC  
ATAAAGTGGTGGAGATAAAATCAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT  
GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA  
GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT  
TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC  
AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA  
GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAACAAGAAGCTACATTTTTTGGCCTTAA  
GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC  
AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA  
ATGTGTTTACTCTCTTTCCTTCCCACATTCTCAATTAAGGTGAGCTAAGCCTCCTCGGTG  
TTTCTGATTAACAGTAAATCCTAAATTCAAAGTGTAAATGACATTTTTTATTTTTATGTCTC  
TCCTTAACATATGAGACACATCTTGTTTTACTGAATTTCTTCAATATTCAGGTGATAGATT  
TTTGTCG







**FIGURE 283**

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC  
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGA**ATGC**  
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA  
GTGTCCTGGGTGAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC  
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG  
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGCG  
GCAAACCTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCAGCTCCT  
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAG**TAA**AACCACAGGCTGG  
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA  
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTACCACATTAGCAATTAAACTGAGAAAT  
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT  
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA  
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG  
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT  
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA





**FIGURE 285**

GTC**ATG**CCAGTGCCTGCTCTGTGCCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA  
GCGGCCCCCATGGGCGGCCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG  
GACCCTGCAGCTGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA  
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTCAGC  
CGGGGCCGGGATGCAGCCAGGAACTTCGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA  
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA  
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT  
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC  
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC  
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCC**TGA**ATCTGCCTGGATGGAAGTGA  
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG  
CCTGTTCACTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC  
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA  
CCCTTTCATGCCTACACACCCCTCATTAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAA

286/330

**FIGURE 286**

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK  
ARNSLGLYGRRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK  
VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ  
IQERLHTAALPA

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK  
ARNSLGLYGRRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK  
VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ  
IQERLHTAALPA



288/330

**FIGURE 288**

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHKKRRVRDKDGLKTQIEKLWT  
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ  
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSSAQGK  
WSDEACRSSKRYICEFTIPK

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHKKRRVRDKDGLKTQIEKLWT  
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ  
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSSAQGK  
WSDEACRSSKRYICEFTIPK











CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGAC**ATGAG**  
GAGGCTCCTCCTGGTCAACAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAG  
CACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAG  
GCCTGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGTCTGTTCCC  
TGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTCAGGGCAGGGGCCCCA  
TCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCTTGAGTCCC  
GAGCCCGACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCC  
CCGGTTGTGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACA  
TCTACCACCCCCAG**TAG**GGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGCCCAGG  
CTGTTGGGACTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCAGCAGGCAAA  
AAAAAAAAAAAAAAAAA

**FIGURE 294**

MRRLLLVTSLVVLLWEAGAVPAKVPKIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL  
FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE  
RPRLWVMPNHQVLLGPEEDQDHIYHPQ

**FIGURE 295**

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTG  
TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATGA**ACCAACTCAGCTTCCTGC  
TGTTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAA  
TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC  
TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCT  
GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG  
CGTGGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCC  
AGAGGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG  
ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG  
CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC  
TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT  
ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGC  
GACGCCCAGAAAACAGCATCTTATTACTCACCCCTATGGCCAGCGGGAATTCAGTGCGGGATT  
TGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG  
TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT  
CCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATATGGAACATCATGTTGGTTA  
CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTTTGTG  
GGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACCTTACCCA  
GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

[illegible]

**FIGURE 296**

MNQLSFLFLFIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN  
 GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG  
 SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGI  
 YQKYPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYSPYQOREFTAGFVQFRVFNNERAAN  
 ALCAGMRVTGCNTEHHCIGGGGYFPEAS PQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL  
 FYR

**FIGURE 297**

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC  
CACGAGGCTGCCGCATCCTGCCCTCGGAACAATGGGACTCGGCGCGCGAGGTGCTTGGGCCG  
CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCGCCCATGAAAGCGCAGCC  
ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAAC  
AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAAC  
CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG  
GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC  
TACACCCAAAACAACAAGTGTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG  
TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT  
TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC  
GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC  
GGTATCGAACCATAGATGAACATGATGCCATCATTTTAAGGAATCCATGGACCAAGGATGGA  
ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAACAATATTCT  
CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA  
AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT  
TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTTGGCT  
GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA  
TGCCATCTGGGCATACAAATAAGAAGTTTGTACAGCACTCAGGATTTTGGGTATCTTTTGT  
AGCTCACATAAAGAAGTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA  
CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG  
TGCTTTAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

**FIGURE 298**

MGLGARGAWAALLLGT LQVLALLGAAHESAAMAASANIENSG LPHNSSANSTETLQHVPSDH  
TNETSNSTVKPPTSVASDSSNTT VTTMKPTAASN TTTPGMVSTNMTSTTLKSTPKTTSVSQN  
TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSILYIG  
CKMYYSRRGIRYRTIDEHDAII

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG  
TCGCGGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCAGCGATGGCGACCCTGTGGGGAGGC  
CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTCGTGCCTGGCGCTTTCGGTGCTGCTGCTGGC  
GCAGCTGTCAGACGCCGCCAAGAATTTGAGGATGTCAGATGTAAATGTATCTGCCCTCCCT  
ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT  
CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA  
ATGCAAATATGAAGAAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA  
TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCATACTGAAGAGG  
CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT  
TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG  
AATATGCACAGCAGCGCTGGAAGCTTCAAGTCCAAGAGCAGCGAAAGTCTGTCTTTGACCGG  
CATGTTGTCTCCTCAGCTAATTGGGAATTGAATTCAGGTGACTAGAAAGAAACAGGCAGACAA  
CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTCATTTTAATACCTTGTTGATTTACCAACT  
GTTGCTGGAAGATTCAAACTGGAAGCAAAACTTGCTTGATTTTTTTTTCTTGTTAACGTA  
ATAATAGAGACATTTTTTAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCCTATTTG  
TGACTTTTACTAATAAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA  
AGCACTCTCTTTTTTCACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTG  
TTGTTGTTGTTTTTTGTTTGTGTTTGGTGGGAGAGGGGAGGGATGCCTGGGAAGTGGTT  
AACAACTTTTTTCAAGTCACTTTACTAAACAACTTTTGTAATAGACCTTACCTTCTATTT  
TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG  
ACTTTTGCCTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT  
CTAAAATGCCTGGTGGCTTTTCACAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG  
CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG  
GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGACACT  
TGCAATAAAGAAATTTTATTTTAAACCAAGCCTCCCTGGATTGATAATATATACACATTTG  
TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT  
AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT  
TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTGAAACATCAA

MATLWGGLRLGSLLSLSCLALSLLLLAQLSDAAKNFEDVRCKCICPPYKENS<sup>GH</sup>IYNKNIS  
QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTII<sup>IY</sup>LSILGLLLLYMVYLTL  
VEPILKRR<sup>LF</sup>GH<sup>AQ</sup>LIQSDDDIGDHQPFANAHDVLARSRSRANVLNKVEYAQQRWKLQVQE<sup>Q</sup>  
RKS<sup>VF</sup>DRH<sup>V</sup>LS



FIGURE 301

[illegible]

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQT  
PGARFQRSHLAEAFAKAKSGGGAGGGSGRGLMGQIIPIYGFGLFYILYILFKVSRIILI  
ILHQ

**FIGURE 303**

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT  
GGATTTGAAAGTTGAGAGCAGCATGTTTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC  
TGGATTATTCCCTGGGCCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG  
GGTGATTGAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT  
AGACTGGACTCTGTACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA  
ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC  
AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA  
AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG  
AGGAGCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTGAGATGGGATGTGTTTTCCAG  
AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA  
GGAGATTGTATTTTCGTTACTACCACAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG  
GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCATCATG  
CTTCAAGGAGTGAGGGAGTCAGATGGAGGAACTACACCTGCAGTATCCACCTAGGGAACCT  
GGTGTTCAAGAAAACCATTTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC  
CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC  
TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA  
GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG  
AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA  
CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA  
CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACCTCACTTGAAAAAAGTCAGGTGGGG  
GAATGCCAAAAACACAGCAAGCCTTTTTGAGAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG  
TGGAGACTCTCTCCTGTGTGTGTCTCTGGGCCACTCTACCAGTGATTTTCAGACTCCCGCTCTC  
CCAGCTGTCCCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG  
CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC  
CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCGTT  
GGATCAGACCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA  
AAAACCAACCCAAATCAA

[illegible][illegible]

**FIGURE 305**

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG  
GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG  
AAACTACATTTTGC AAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA  
GGATGAAGATGGATACATCACCTTAAATATTAAAACCTCGGAAACCAGCTCTCGTCTCCGTTG  
GCCCTGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG  
ATGGTTGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA  
TGAGAATGAAAATCGCACAGGAACCTTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG  
TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCCTGTGACACAAAC  
TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACCTTAACATGGGAAGAGAG  
TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG  
AGTACATCAAAGCCAGGACTCATTTAATTTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT  
GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGAAGA  
TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG  
AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT  
**TAA**TGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT  
ATGAATGCATCAGTAGCTGAAAAAAAAAAAAA

306/330

**FIGURE 306**

MQDEEDGYITLNIKTRKPALVSVGPASSSWWRVMALILLILCVGMVVGLVALGIWSVMQRNYL  
QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE  
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEEFL  
EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE  
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEEFL  
EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

**FIGURE 307**

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG  
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCCGCGCGGGAGCCGGACCGC  
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA  
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG  
CGGAGGAGAAGGAGGAGGAGGCGAACCAGAGAGGGGCAGCAAAGAAGCGGTGGTGGTGGG  
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG  
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC  
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG  
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC  
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT  
CTGTTTAAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT  
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACCTTTTACACCTGAGTGCA  
AATTCAAAGAATCAGTGTTTGAATAATTATTATGTGACATATTCATCAATGATATACCGTCAG  
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA  
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT  
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC  
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC  
**G****TAG**CCAGTGAGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT  
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTGAGTGACATTTACCAAACAAACAGG  
CAGAGTTCACCTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

308/330

## **FIGURE 308**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLVFNRVCLFGSKKRRRRRP  
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTDKEDSTYTLFNLI PVGLRVVAIQGVQTKLYLA  
MNSEGILYTSLEFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLENKEGEIMKGNHVK  
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMHNEST

### **N-glycosylation site.**

amino acids 242-246

### **Glycosaminoglycan attachment site.**

amino acids 165-169, 218-222

### **Tyrosine kinase phosphorylation site.**

amino acids 93-100

### **N-myristoylation site.**

amino acids 87-93, 231-237

### **ATP/GTP-binding site motif A (P-loop).**

amino acids 231-239

### **HBGF/FGF family proteins**

amino acids 78-94, 102-153

seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498  
subunit 1 of 1, 245 aa, 1 stop  
MW: 27564, pI: 10.18, NX(S/T): 1  
MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLVFNRVCLFGSKKRRRRRP  
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTDKEDSTYTLFNLI PVGLRVVAIQGVQTKLYLA  
MNSEGILYTSLEFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLENKEGEIMKGNHVK  
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMHNEST  
N-glycosylation site.  
amino acids 242-246  
Glycosaminoglycan attachment site.  
amino acids 165-169, 218-222  
Tyrosine kinase phosphorylation site.  
amino acids 93-100  
N-myristoylation site.  
amino acids 87-93, 231-237  
ATP/GTP-binding site motif A (P-loop).  
amino acids 231-239  
HBGF/FGF family proteins  
amino acids 78-94, 102-153



**FIGURE 309**

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG  
 ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG  
 CAACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCT  
 CCATTTAGAGTGTAGCAAAGGAAAAAACACCAAGGTTGGGTCCTTCCTGACATTGGCAGTG  
 CCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC  
 AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA  
 ACCACGTCTTGAAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTG  
 GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAAACCTCGAGGGTGGGATCCACTGAGG  
 AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCACGGGTGGTAACCTGGCTGCT  
 GTGGAGGGGGGTACGTGAGGGGGGGGTCTGGGGCTTATCCTCAGGTCCTGTGGGTGGGGCAG  
 CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG  
 CGCGCTCCGGGCGCCTGCCGGTTTGGGGGTGTCTCCTCCCGGGGCGCT**ATG**GCGGCGCTGGC  
 CAGTAGCCTGATCCGGCAGAAGCGGGAGGTCCGCGAGCCCGGGGGCAGCCGGCCGGTGTGCG  
 CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG  
 CTGTCCAAGGTGCGACTGTGCGGGGGGCGGGCCCGCGCGGCCGGACCGCGGCCCGGAGCCTCA  
 GCTCAAAGGCATCGTCACCAAACCTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG  
 ACGGAAGCATCCAGGGCACCCCAGAGGATAACAGCTCCTTCACCCACTTCAACCTGATCCCT  
 GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC  
 TGAGGGACTGCTCTACAGTTCGCCGCATTTACAGCTGAGTGTGCTTTAAGGAGTGTGTCT  
 TTGAGAATTACTACGTCTGTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGGCC  
 TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA  
 GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC  
 ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCC**TGA**AATGTAGTCCCTGGACTG  
 GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACAACCTGT

**FIGURE 310**

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR  
GPEPQLKGIIVTKLFCRQGGFYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY  
MAMNAEGLLYSSPHFTAECRFKECVFENYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR  
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

**Tyrosine kinase phosphorylation site:**

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

**FIGURE 311**

**ATG**GCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG  
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACG  
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG  
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT  
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCT  
TCAACCTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT  
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTT  
TAAAGAATCTGTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGG  
AATCTGGTAGAGCCTGGTTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA  
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG  
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA  
GCACAAGTGCCTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

**FIGURE 312**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDPRPSASRRRSSPSKNRGLCNGNLVDIFSKVRIFGLKKRRLR  
RQDPQLKGIVTRLYCRQGYLQMHDPGALDGTKDDSTNSTLFLNLI PVGLRVVAIQGVKTGLY  
IAMNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR  
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKSITT

**N-glycosylation site.**

amino acids 100-104, 242-246

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 28-32, 29-33

**Tyrosine kinase phosphorylation site.**

amino acids 199-207

**N-myristoylation site.**

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

**HBGF/FGF family proteins.**

amino acids 104-155, 171-198

**FIGURE 313**

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAAT  
 GAAGGATGCAGGACGCAGCTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGA  
 ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG  
 AAATAAACAGAGTTAGACCCGCGGGGGTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCC  
 CTCCCCACCCCCAAAAAAAGGATGATTGGAATGAAGAACCAGGATTCACAAAGAAAAAAGTATGTTCAATTT  
 TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAA  
 AGAACTGGTGTGGTGGTGTTCCTTTCTTTTTGAATTTCCCACAAGAGGAGAGGAAATTAATAATACATCTGC  
 AAAGAAATTTAGAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAAACCAGCAGAGCA  
 CAGTTGGATTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT  
 TTTTTAAATTTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTGTTCTTAACCACCTGGATTTCCATCT  
 GGATGTTGCTGTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATG  
 TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCT  
 GCTTGTGGTGTCTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGT  
 GCTCCTGCAGCAACCAAGTTCAGCAAGGTGATTTGTGTTTCGGAACCACTGCGTGAGGTTCCGGATGGCATCTCC  
 ACCAACACACGGCTGCTGAACCTCCATGAGAACCATAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAG  
 GCACTTGGAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTTGGGGCTTTCAATGGTCTGGCGA  
 ACCTCAACACTCTGGAACCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAA  
 CTGAAGGAGCTCTGGTTGCGAAACAACCCCATGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCTTCTTT  
 GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCA  
 TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAACTAGATGAG  
 CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTCCAGGGTTTGATGCACCTTCAAAA  
 GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA  
 ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGATCATCTAGAGCGGATACAT  
 TTACATCACAACCCTTGGAACGTAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC  
 GAACACAGCTTGTTGTGCCCCGGTGTAACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA  
 ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT  
 GAGCTGAAATGTGCGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA  
 TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTACAAATGTAACGTGTGCAAGATA  
 CAGGCATGTACACATGTATGGTGAATAATCCGTTGGGAATACTACTGCTTCAGCCACCTGAATGTTACTGCA  
 GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG  
 GACCACAGATAACAATGTGGGTCCCACTCCAGTGGTGCAGTGGGAGACCACCAATGTGACCACCTCTCTCACAC  
 CACAGAGCACAAAGTCGACAGAGAAAACCTTACCATCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT  
 GATGAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT  
 GGTCAATTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCACGCCCCAACAGGACTGTTGAAATTA  
 TTAATGTGGATGATGAGATTACGGGAGACACACCATGGAAGCCACCTGCCCATGCCTGCTATCGAGCATGAG  
 CACCTAAATCACTATAACTCATACAAATCTCCCTTCAACCACACAACAACAGTTAACACAATAAATTCATACA  
 CAGTTCAGTGATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATCTAAACA  
 TTTACAGAGTTACAAAAACAAACAATCAAAAAAAGACAGTTTATTAATAATGACACAAATGACTGGGCTAA  
 ATCTACTGTTTCAAAAAGTGTCTTACAAAAAACAAGAAAGAAATTTATTTATTAATAATCTATTG  
 TGATCTAAAGCAGACAAAA

**FIGURE 314**

MLNKMTLHPQQIMIGPRFNRALFDPELLVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC  
 VRKNLREVPDGI STNTRLLNLHENQIQI IKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA  
 NLNTLELFDNRLTTIPNGAFVYLSKLELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS  
 YISEGAFEGLSNRLRYNLAMCNLREIPNLTPLIKLDLDELDSGNHLSAIRPGSFQGLMHLQKL  
 WMIQSQIQVIERNADFNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHNPWNCNDIL  
 WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE  
 LKCRASSTLSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNTVTQDTGMYTCMVNSVGN  
 TTASATLNVTAATTTTFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPO  
 STRSTEKTFTIPVTDINSIGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQN  
 HHAPTRTVEIINVDDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS  
 VHEPLLIRMNSKDNVQETQI

**Signal sequence:**

amino acids 1-44

**Transmembrane domain:**

amino acids 523-543

**N-glycosylation site.**

amino acids 278-282, 364-368, 390-394, 412-416, 415-419,  
 434-438, 442-446, 488-492, 606-610

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

**Casein kinase II phosphorylation site.**

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

**N-myristoylation site.**

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,  
 391-397, 422-428, 433-439, 531-537

GCGCCGGGGAGCCCATCTGCCCCCAGGGGGCACGGGGCGCGGGGGCCGGCTCCCCGCCGGGCACAT  
 GGCTGCAGCCACCTCGCGCGCACCCCCGAGGCGCCGCGCCCAGCTCGCCCGAGGTCCGTGCGA  
 GCGCCCCGGCCGCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC  
 GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA  
 CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACCTTGGCCCTGCCACCATCAACTGGGGC  
 TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCATAATGAAGGGAACCAAAAA  
 GTGGTGATCACTTACTCCAGTCGTATGTCTACAATAACTGACTGAGGAACAGAAGGGGCC  
 AGTGGCCTTTGCTTCCAATTTCTTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGC  
 CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT  
 GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC  
 AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT  
 ACTGGCAGCGAATCCGAGAGAAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT  
 GACTACAACCACCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTACTCTGGACTGTA  
 CCAGTGACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT  
 ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG  
 ATTTTCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA  
 GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAGCCCGTCTTGTGAAACCCAGCT  
 CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCTCCACTCGCTCCACAGCAAAT  
 AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC  
 CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG  
 CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA  
 ACGGT**CTGA**ATTACAATGGACTTGACTCCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC  
 TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA  
 GTGAGCATTGCACGGAACAGATTAGATGAGCATTTTCTTTATACAATACCAAACAAGCAAA  
 AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG  
 AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG  
 AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT  
 TTCAAGAGGAAATGGGATGCTGTTTGTAATTTTCTATGCATTTCTGCAAACCTATTGGATT  
 ATTAGTTATTAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC  
 TGAGCTAACCACCTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC  
 TTCATTTGTATAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA  
 AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCCAAATAAC  
 TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT  
 ATGAGGATTGTTGACAAACATTTAGAAATATATAATGGAGCAATTTGTGGATTTCCCTCAAAT  
 CAGATGGCCTCTAAGGACCTTCTCTGCTAGATATTTCTGGAAGGAGAAAAACAACATGTCAATT  
 TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA  
 CCAACATACCATTTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG  
 GTGGACTAGAAAGGGAGATTAGATCAGTTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA  
 TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC  
 ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

316/330

## **FIGURE 316**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTGLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV  
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV  
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID  
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI  
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS  
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 232-251

MSLLLLLLLLVSYYVGTGLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV  
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV  
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID  
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI  
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS  
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV



**FIGURE 317**

[illegible]

**FIGURE 318**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA  
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTLNLQLSDIGTYQCKVKK  
APGVANKKIHLVVLVKPSGARCIVDGSEEIGSDFKIKCEPKESLPLQYEWQKLSDSQKMPT  
SWLAEMTSSVISVKNASSEYSGTYSTVRNRVGSQCLLRNLNVVPPSNKAGLIAGAIIGTLL  
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM  
EGYSKTQYNQVPSSEDFERTPQSPTLPPAKFKYPYKTDGITVV

**Signal sequence.**

amino acids 1-19

**Transmembrane domain:**

amino acids 236-257

**N-glycosylation sites.**

amino acids 106-110, 201-205, 298-302

**Tyrosine kinase phosphorylation sites.**

amino acids 31-39, 78-85, 262-270

**N-myristoylation sites.**

amino acids 116-122, 208-214, 219-225, 237-243, 241-247,  
245-251, 296-302

**Myelin P0 protein.**

amino acids 96-125

TGAAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGA  
ATGGTGAAGGTGCCTGTCTAACTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCC  
CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTTCG  
CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG  
AGAGATGAATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC  
CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG  
CTTCTCTGGAGCCAGGTATCAGGGGCCAGGGCCAAGAATTCACCTTTGGGCCCTGCCAAGT  
GAAGGGGGTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG  
CTCAGGATAACATCACGAGTGCCCCGGCTGCTGCAGCAGGAGTTCTGCAGAACGTCTCGGAT  
GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAAGTGTTTTCAAAAA  
CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAAC  
TTGTTCTCATCGTGTCACAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC  
AGTGCACACAGGCGGTTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC  
TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC  
TCTGAATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTTGTTCCCTGTGTCAATTTCA  
AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATT  
TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG  
GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACCTCTATTT  
AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG  
CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG  
TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTGGCCTGTCTTTGGATT  
GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACT  
GTCTTCTGATATCTGCAGGGACAGAGCATTGGGGTGGGGGTAAAGGTGCATCTGTTTGAAAAG  
TAAACGATAAAATGTGGATTAAAGTGCCCAGCACAAAGCAGATCCTCAATAAACATTTTCATT  
TCCCACCCACACTCGCCAGCTCACCCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTT  
TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC  
ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCTTGCTAATAAAAGACAACATAA  
CTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 320**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRQLQSLWTLARPFPCPLLATASQMOMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK  
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNH  
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRREFLLFRRAFKQLDVEAAL  
TKALGEVDILLTWMQKFYKL

**Signal sequence:**

amino acids 1-42

**N-glycosylation sites.**

amino acids 85-89, 99-103, 126-130

seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002  
subunit 1 of 1, 206 aa, 1 stop  
MW: 23799, pI: 9.12, NX(S/T): 3  
MNFQQRQLQSLWTLARPFPCPLLATASQMOMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK  
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNH  
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRREFLLFRRAFKQLDVEAAL  
TKALGEVDILLTWMQKFYKL  
Signal sequence:  
amino acids 1-42  
N-glycosylation sites:  
amino acids 85-89, 99-103, 126-130

**FIGURE 321**

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGC**ATG**AAGTTACAGTGTGTTTCCCTTTGGCTC  
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTT  
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAGAGCCATCCAAGCTAAGG  
ACACCTTCCCAAATGTCACTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAAGCCCTTA  
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA  
TCAGGAGCCAAACCCCAAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA  
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC  
AATGCCACCAGAGTCATCCATGACAACTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA  
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT  
CAGCT**TGA**TGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT  
GTGGGAGACAGCCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC  
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGCAAAAAGTCTACTGTGGTAT  
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT  
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT  
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

**FIGURE 322**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST  
 LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ  
 RQCHCRQEATNATRVIHNDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

**Signal sequence:**

amino acids 1-18

**N-glycosylation sites.**

amino acids 56-60, 135-139

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 102-106

**N-myristoylation site.**

amino acids 24-30

**Actinin-type actin-binding domain signature 1.**

amino acids 159-169

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

**FIGURE 323**

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG  
AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT  
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC  
GGTTCTATCGATAATCTCAGCACCAAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCCGCCT  
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA  
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC  
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC  
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA  
TGAGCAGAAGATACCTCTGCATGGATTTTCCAGAGGCAACATTTTTGGATCACACTATTTTCGAC  
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC  
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC  
CACCCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACCTCAACACC  
CCCATACCACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT  
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTTACAGGAGCTCCCGAGCG  
CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAAC  
ACGCACGCTGGGGGAACGGGCCCCGGAAGGCTGCCGCCCTTCGCCAAGTTCATC**CTAG**GGTCG  
CTGG

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG  
AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT  
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC  
GGTTCTATCGATAATCTCAGCACCAAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCCGCCT  
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA  
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC  
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC  
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA  
TGAGCAGAAGATACCTCTGCATGGATTTTCCAGAGGCAACATTTTTGGATCACACTATTTTCGAC  
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC  
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC  
CACCCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACCTCAACACC  
CCCATACCACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT  
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTTACAGGAGCTCCCGAGCG  
CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAAC  
ACGCACGCTGGGGGAACGGGCCCCGGAAGGCTGCCGCCCTTCGCCAAGTTCATC**CTAG**GGTCG  
CTGG

**FIGURE 324**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWGGLIHLYTATARNSYHLQIHKNGHVD  
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLNGY  
DVYHSPQYHFLVSLGRAKRAFLPGMNPPPYSQLSRNEIPLIHFNTPIPRRHTRSAEDDSE  
RDPLNVLKPRARMTAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFA  
KFI

**Important features of the protein:**

**Signal peptide:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 175-179

**N-myristoylation site.**

amino acids 33-39, 100-106, 225-231, 229-235

**HBGF/FGF family proteins**

amino acids 73-124

seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238  
subunit 1 of 1, 251 aa, 1 stop  
MW: 27954, pI: 9.22, NX(S/T): 1  
MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWGGLIHLYTATARNSYHLQIHKNGHVD  
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLNGY  
DVYHSPQYHFLVSLGRAKRAFLPGMNPPPYSQLSRNEIPLIHFNTPIPRRHTRSAEDDSE  
RDPLNVLKPRARMTAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFA  
KFI  
Important features of the protein:  
Signal peptide:  
amino acids 1-24  
cAMP- and cGMP-dependent protein kinase phosphorylation site.  
amino acids 175-179  
N-myristoylation site.  
amino acids 33-39, 100-106, 225-231, 229-235  
HBGF/FGF family proteins  
amino acids 73-124



**FIGURE 325**

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**  
 GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG  
 GAGCTCTGCAGGCCCCAGCACCCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG  
 TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG  
 ACCTCTTCTAGGGCCTCAACCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA  
 GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC  
 TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC  
 ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA  
 TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA  
 CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCCTCTCCGACGGCCCCCATCCAGTCATC  
 ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC  
 GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT  
 GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG  
 GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACCTTCCAGCATCCCTGGGGCCTCAGA  
 CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC  
 CTGACTCCACTGAAGCAAACACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC  
 ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA  
 CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA  
 CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC  
 AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTGAGCAGTGGGCAAAACAACCTTC  
 CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC  
 CTTAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT  
 CTTCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA  
 GATCACAACCTCAGCGAAGACCAGATGAAGCCCCAACAGCCACGCCCACGACTGCCCCGAC  
 GAGGCCGACCACAGACG**TGA**GTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG  
 TGGCTTCCCCGGAAGACCTCACTGACCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC  
 CGGGAACCTCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA  
 ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCCTAGCCTG  
 GGCCCCCACCAGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG  
 AAGGGCAGCATGTCCAAGCCCCAATCCCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC  
 CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCCAGAGGTGTCCTTGGAATCACCTTGG  
 CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA  
 TTAAAATTCACTCAGTGTGGCCCCAAAAAAA

**FIGURE 326**

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTLA  
 ETSSRASTPAGPIPEAETRGAKRISPARETRSF TKTSPNFMVLIATSVETSAASGSPEGAGM  
 TTVQTITGSDPEEAI FDTLCTDDSS EAKTLTMDILT LAHTSTEAKGLSSESSASSD GPHPV  
 ITPSRASESSASSD GPHPVITPSRASESSASSD GPHPVITPSWSPGSDVTLLAEALVTVTNI  
 EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL  
 STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTL SGALVTVSRNPLEETSALSVETPSY  
 VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNF TPSETPTMDIATKGPFPTSRD  
 PLPSVPPTTTNSSRG TNSTLAKITTS AKTTMKPQQPRPRLPGRGRPQT

**N-glycosylation sites:**

amino acids 252-256, 445-449, 451-455

**cAMP-and cGMP-dependent protein kinase phosphorylation site.**

amino acids 84-90

**Casein kinase II phosphorylation sites.**

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,  
 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,  
 404-408, 414-418, 431-435

**N-myristoylation sites.**

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-  
 320, 349-355, 386-392, 397-403, 449-455

**ATP/GTP-binding site motif A (P-loop).**

amino acids 385-393

**FIGURE 327**

GCGGAGCATCCGCTGCGGTCTCGCCGAGACCCCGCGCGGATTCGCGCGGTCTTCCCGCGG  
 GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCTCTCGACGCCA  
 GAGAGAAATCTCATCATCTGTGCAGCCTTCTTAAAGCAAATAAGACCAGAGGGAGGATTAT  
 CCTTGACCTTTGAAGACCAAACTAACTGAAATTTAAAATGTTCTTCGGGGGAGAAGGGAG  
 CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT  
 GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTCTAAGGGAATC  
 AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC  
 AAAAAACATATCAGGGGACAAAGCATGTAACCTTGATGATCTTCGACACTCGAAAAACAGCTA  
 GACAACCCAACCTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA  
 AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG  
 CCAAGAGTTACCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC  
 TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT  
 CAGAAGTTTGGATCCTCAGATCACCTGGAGAACTATTTAAGATGGATGAAGCAAGTGCCCA  
 GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCTCTGATCAAGAAA  
 TAGCTCATCTGCTGCCTGAAAATGTGAGTGCGCTCCAGCTACGGTGGCAGTTGCTTCTCCA  
 CATAACACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC  
 TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC  
 AGCCTCCCACGACCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG  
 GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA  
 AACCATAACGTTTACAGAAATCTCCAACCTAACTTTGAACACAGGGAATGTGTATAACCCTA  
 CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAACTGCTTCTCGGGAAGGT  
 AGGGAGGCCAGTCCAGGCAGTTCTCCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC  
 ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGCTCTGTTCTGGTGATAGGCC  
 TCGTCTCTCTGGGTAGAATCCTTTCGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT  
 TATTTGATCAATGGGATCTATGTGGACATCTAAGGATGGAACCTCGGTGTCTCTTAATTCATT  
 TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG  
 TATTTTGAAGACAGGAAAATGCCCCCTTCTGCTTTCTTTTTTTTTTTTGGAGACAGAGTCTT  
 GCTCTGTTGCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCAGCAACCTCCGTCTC  
 CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA  
 CCACACCTGGGTGATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTG  
 GTCTCAAACCTCTGACCTAGTGATCCACCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG  
 CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTGAAGAAGGAATGAAGTG  
 GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAACAAAGCTCT  
 ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAATTCAACTGGCTTTTATGCAAA  
 GAAACAGGTTAGGACATCTAGGTTCCAATTCATTACATTCTTGGTTCCAGATAAAATCAAC  
 TGTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTATATGGATTCTTTTAAACTTATT  
 CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTACTCAA

**FIGURE 328**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLKGIKGNPVTSTQED  
CINCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP  
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF  
KMDEASAQLLAYKEKGHSQSSQFSSDQEIHLLENVVSALPATVAVASPHSTTSATPKPATLL  
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP  
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV  
PENQYGLPFEKWLIGSLLFGVFLVIGLVLLGRILSESLRRKRYSLDYLINGIYVDI

**Signal sequence.**

amino acids 1-25

**Transmembrane domain.**

amino acids 384-405

**N-glycosylation sites.**

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 415-419

**Tyrosine kinase phosphorylation site.**

amino acids 50-57

**N-myristoylation sites.**

amino acids 4-10, 48-54, 315-321

**FIGURE 329**

CTCCTCCACCGGTGTCCAGACGCCCAGAAATGCGGCTTCTGTGCTCTGCTATGGGGTTGCCTGCTGCT  
 CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTTGAAGGGGACACTGTGT  
 CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT  
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT  
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA  
 ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG  
 TCTTTACTGATCTCTCTGTTTCGTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCCTTCTCCAC  
 CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAGCTCAGCAAACCCAGCCC  
 CAGGATTGACTTCTCTGCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG  
 GCTGAGGCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC  
 AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC  
 TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG  
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCCTGGTGCTGCTGAGCCTTCTGTGAGC  
 CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA  
 CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC  
 CTTTCCCAGGCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA  
 GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT  
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG  
 TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGATGTTCCA  
 GCCTGACCTAGAAGCGTTTGTAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG  
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA  
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCCTGGGC  
 CTCATGCCCAGTGTGCGACCCTGCCTTCTCCCTCCACTCCAGACCCCACCTTGTCTTCCCTCCC  
 TGGCGTCCCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT  
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT  
 GTGAAAAACGTGATTCTGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG  
 GACTCTGAATTCTAACAATGCCCAGTGAAGTGTGCACCTGAGTTTGAGGGCAGTGGGCTG  
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC  
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG  
 TCCAGGCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG  
 TTGCCTTTTNCATTTGCCCTCCCTGGNCCATGCCTTCTTGCTTTTGAAAAAATGATGAAGA  
 AAACCTTGGCTCCTTCCCTTGTCTGAAAAGGGTTACTTGCTATGGGTTCTGGTGGCTAGAGA  
 GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG  
 ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA  
 GCACAATACTATTTTTTTTTCTTTTTTCCATTATTATTGTTTTTTTAAAGACAGAATCTCGTGCT  
 GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCCTGGGTTCAAGTGATT  
 CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT  
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGAC  
 CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG  
 TCTGGCCCTATTTCCCTTTAAAAAGTGAAATTAAGAGTTGTTTCAAGTATGCAAACTTGGAAG  
 ATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT  
 TATTTTCGTTTTTGTGTACTTCTTCCACTCTTTTCTTCTTACATAAATTGCCGGTGTCTCTT  
 TTTACAGAGCAATTATCTTGTATATACAACCTTGTATCCTGCCTTTTCCACCTTATCGTTCC  
 ATCACTTTATTCAGCACTTCTCTGTGTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCA  
 GCTGCATAAAAAAAAAAAAAA

330/330

## **FIGURE 330**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGP EEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS  
GTIYAE EEGQETMKGRVSIRDSRQELSLIVTLWNLT LQDAGEYWCGVEKRGPD E SLLISLFV  
FPGPCCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG  
TSQYGHERTSQYTGTSPHPATSP PAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSI PMVRI  
LAPVLVLLSLLSAAGLIAFC SHLLLWRKEAQQATETQRNEKFWLSRLTAAEEKEAPSQAPEGD  
VISMPPLHTSEEEELGFSKFVSA

### **Important features:**

#### **Signal peptide:**

amino acids 1-17

#### **Transmembrane domain:**

amino acids 248-269

#### **N-glycosylation site.**

amino acids 96-99

#### **Fibrinogen beta and gamma chains C-terminal domain.**

amino acids 104-113

#### **Ig like V-type domain:**

amino acids 13-128